Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTC
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCCTCGCCGCTGTG
CCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACTGGAGTTACTCATGCTGCACAGT
AATGGCATTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAGGTC
TTAAAAATGAGCTATAACAAAGTCCAAATCATTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAACTGCAGATGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAAACATAGAGACTTAATACATCGGCGGCGTGGGGATTCCACG
CTCCGGCGATTCAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAACTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCCTCCAGATGAAGAATTCATGGTTC
TGAAAACTAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTCACACACCACTGCCTCGTCTTTATTTCACATTCCTAGAAACAA
CAATACAGGTAACTTCCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGAGAGTTAAAAACCCACATAGAACCCCAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGCTCACAG
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCCACAGTGCCCTCCCCAAAACTA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAAACTGGGAAAGATTCAGTGGAAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
CAGTGACAAGTCCTACTGCTACTGCATCTGTCATTATGTCTGAAACCCAACGAACAA
GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
TGATACCCCCTTGGCTTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAG
TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCACTCCTCAGC
AGTGGGGCGCCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
AAAGGCCAGCAGTAAGCATGTCCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
TGGTCAAAACCCAACTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
AGACGTTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGTCTGTG
GTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
GGAAGTACTGTGGAACTAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGA
TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
GTGTTCACTGGGTCCTTTATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATT
CCAGATTTTTCTTGTATCCAAATGGAACTCTGTATATAAGAAGCATCGCTCCTTCAGT
GAGGGGCACTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCTCTCAG
AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCATTTCATTCTCCAATGACAGGTTCACATTTC
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAAGCC
ACAGCCATTCGGCACTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTG
GCCAGGCAGTTATTAGCGTGTCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGAA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACTCTTAACTTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCCTCGTGTCTGTTTCCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGCGGTCTGTGTAAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

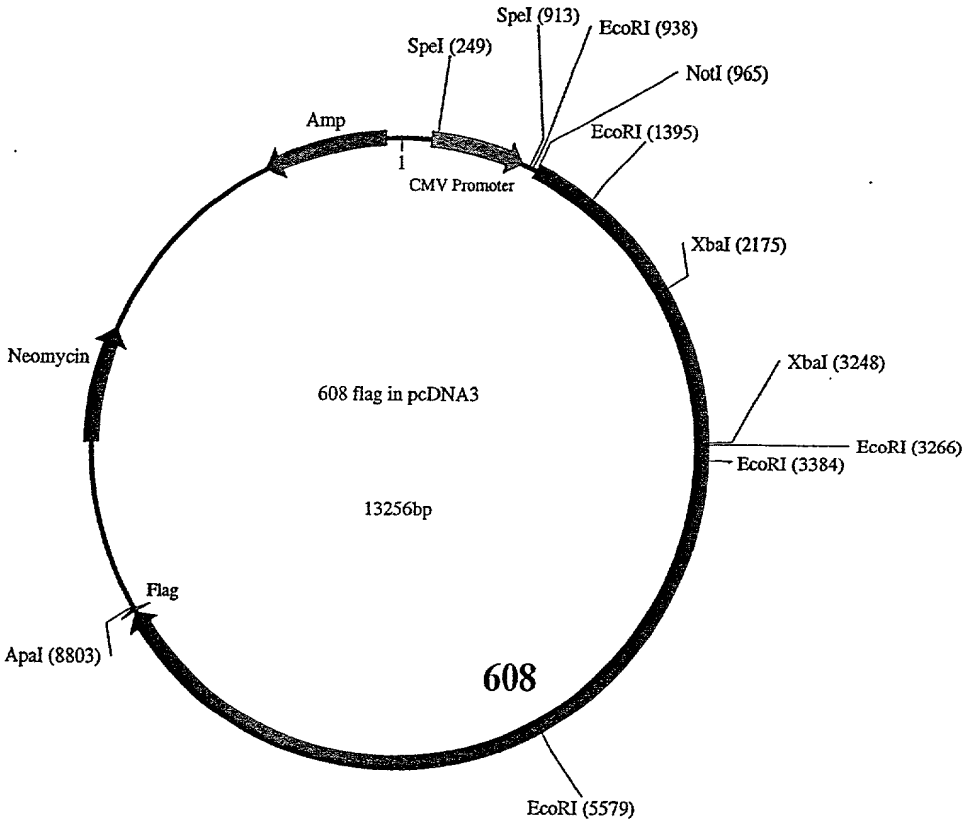

SpeI (913)
EcoRI (938)
SpeI (249)
NotI (965)
Amp
EcoRI (1395)
1
CMV Promoter
XbaI (2175)
Neomycin
XbaI (3248)
608 flag in pcDNA3
EcoRI (3266)
EcoRI (3384)
13256bp
Flag
ApaI (8803)
608
EcoRI (5579)

Figure 3

MQVRGREVSGLLISLTAVCLVVTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNGS
ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
ADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMMNNPK
LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
DSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRK
NKHRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIIFPSVAE
IRDSAQAGRASSQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQIPITGV
SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
KNPHRTPVLRRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGLTVATAALS
VPSSSHSALPKTNNVGVIAEESTTVVKKPLLLFKDKQNVDIEIITTTTKYSGGESNHVIPTE
ASMTSAPTSVSLGKSPVDNSGHLSMPGTIQTGKDSVETTPLPSPLSTPSIPTSTKFSKRKTP
LHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDSTTLLTSPPPALST
TMAATQNKGTEVVSGARSLSAGKKQPFTNSSPVLPSTISKRSNTLNFLSTETPTVTSPTAT
ASVIMSETQRTRSKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTTADTPLAFSHSP
RQDDGGNVSAVAYHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQESTTVKRASDTP
PPLLSSGAPPVPTPSPPPFTKGVVTDSKVTSAFQMTSNRVVTIYESSRHNTDLQQPSAEAS
PNPEIITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSPWPEHKYQLKSYSETIEKGKRPA
VSMSPHLSLPEASTHASHWNTQKHAEKSVFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
GKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV
SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
RPTVSWILANQTVVSETAKGSRKVWVTPDGTLIIYNLSLYDRGFYKCVASNPSGQDSLL
VKIQVITAPPVIIEQKRQAIVGVLGGSLKLPCTAKGTPQPSVHWVLYDGTELKPLQLTHS
RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVVILTVEEGETIPRIETASQKWTEVN
LGEKLLLNCSATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
YLCVARNKMGDDLVLMHVRLRLTPAKIEQKQYFKKQVLHGKDFQVDCKASGSPVPEV
SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNVGMAEEGDYICSAQNTLGK
DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETAVLDCEVTGEPKPNVFWLLPSNNVISFS
NDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLILYKATRNKSGKYRCAARNKVGYIEKLIL
LEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTPGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNYICRAQNSVGQAVISVSVMVVAYPPRIINYLPRNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

Figure 4

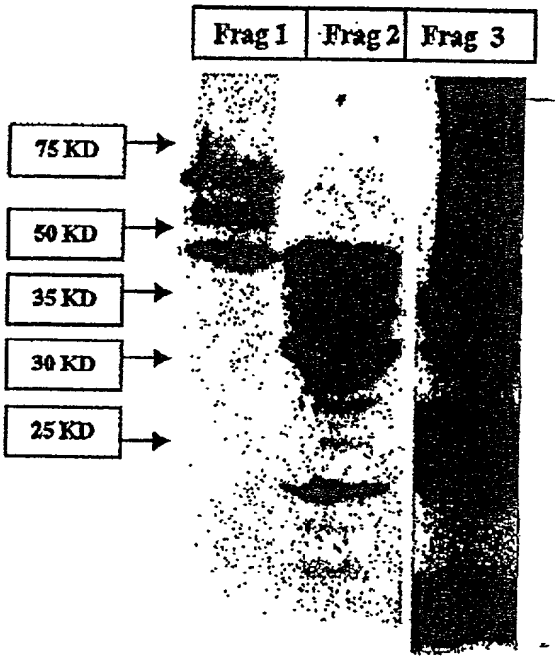


Figure 5

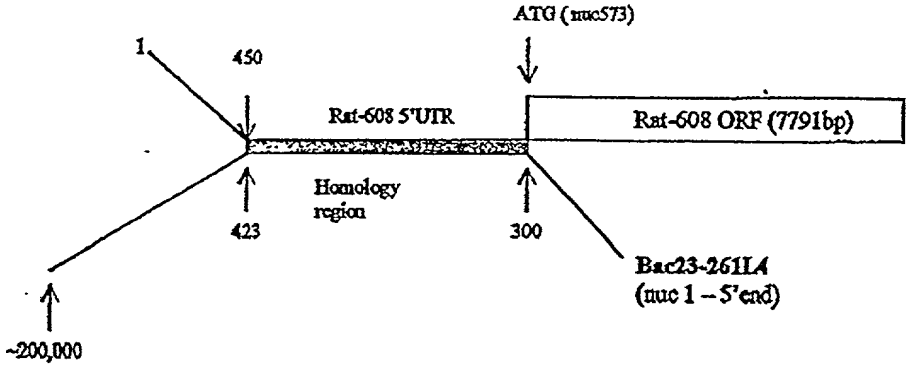


Figure 6

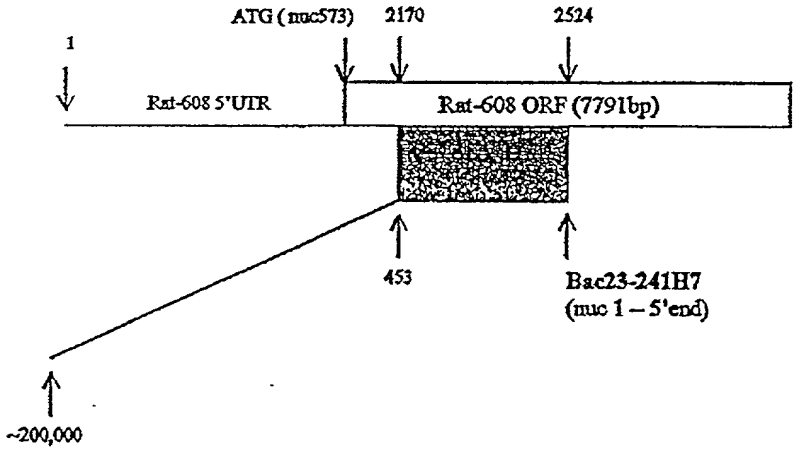
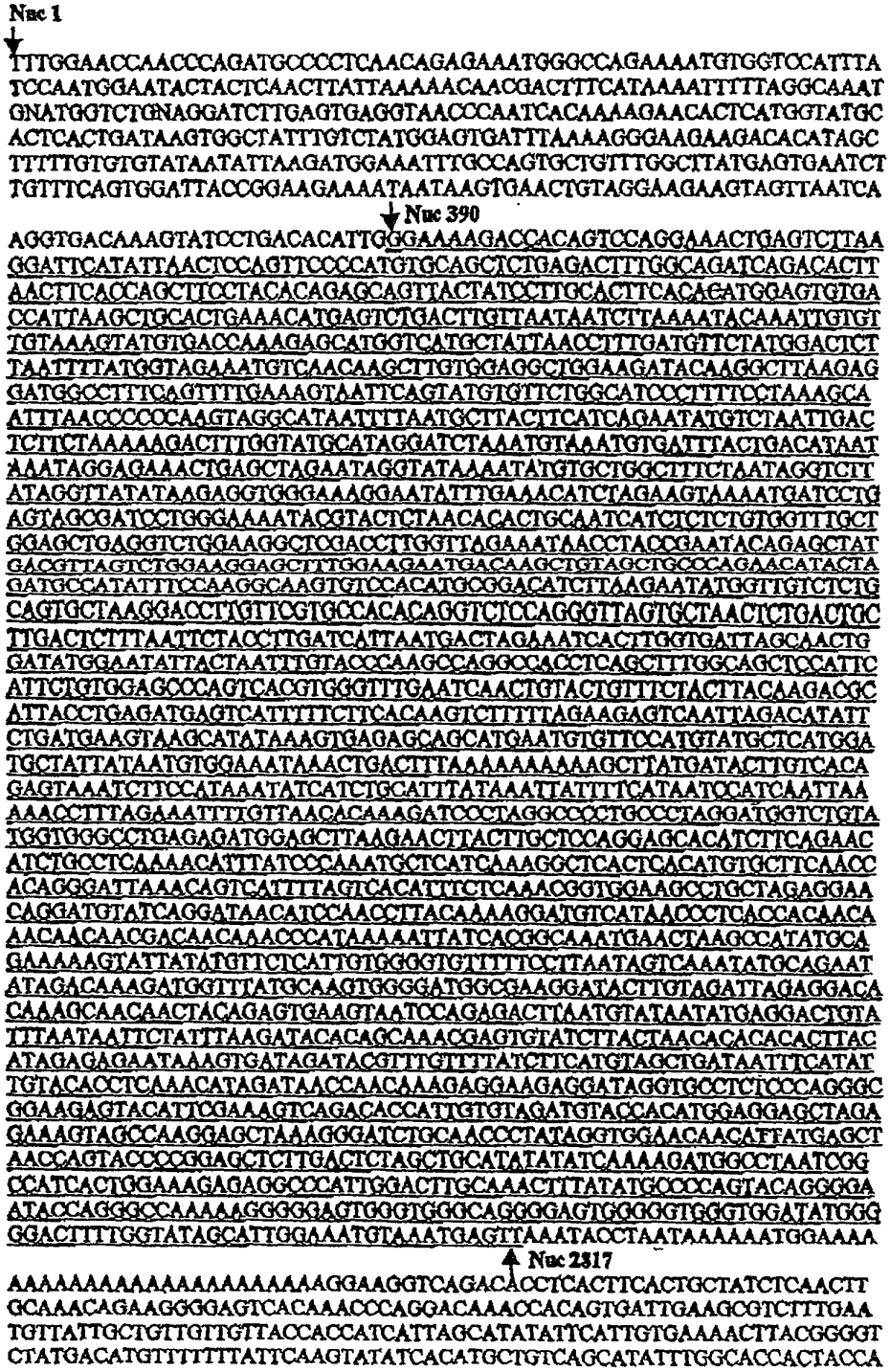
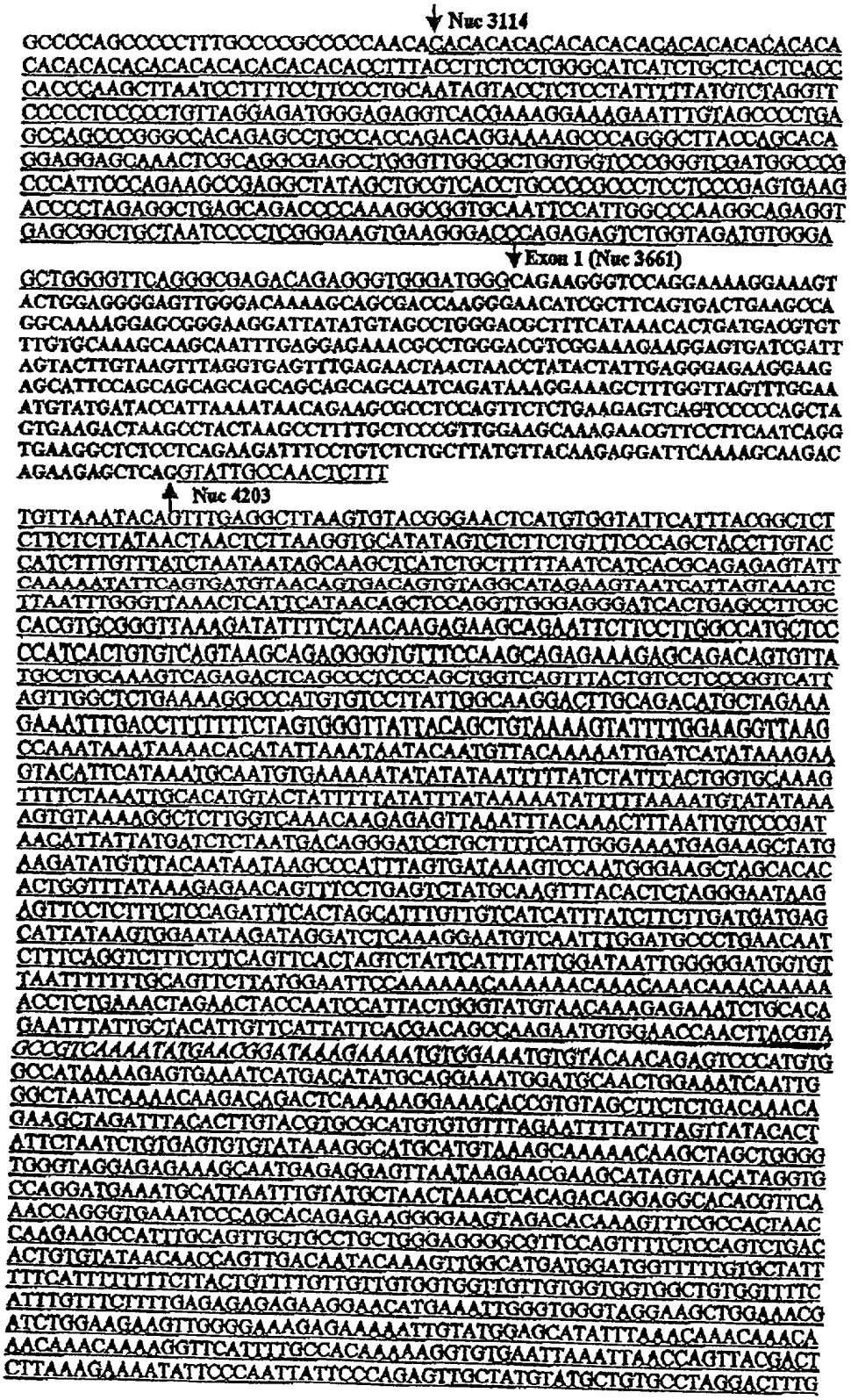


Figure 7




Nuc 3114
Exon 1 (Nuc 3661)
Nuc 4203

Exon 2 (Nuc 6559)

CTTGAACTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG*ATG*CAGAAGAGAGGCAGGGAA
GTCAGCTGCTTGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTC
TGTCCTCGCCGATGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGGACCTGACC
TCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

Nuc 6755

GGACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCCTGGTACTTTGCCCCCC
AAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTCATACAAACCTTATATTG
TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCCCCTCTTAGTT
GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTTAAATGAAAAAGAT
ACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAATTGTCATTGAATATTTTAA
AAAAGCATTTCTTTTGCAATTTATAAATACCCATTACAAAATGGCTTACTTAAAATACTTG
CCTTACTAAATCTGACAAATTATGGTGATATTTTGAAGGTTTATGAAAATTTGTTTATGTGT
ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTGCCATTAGTGAGCATGTACAGT
ATGCCAAACACTATTGTTCACGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT
TATAACCGTATACCCAGTGCCTTGGAAGCGATTGCAAACAGTAAAGACTGACATTGTGTT
CTCCCTATGAGGGAGGGGCCTTGGGCTGAGCACTTTGCAATGAGCATTTGCTCATTGTGCT
GGCAGGTTTTATGATAACTTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACTTCAACT
GAGAACATGCCTGAAGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTTCTTAATTAG
TGATTCATGGGGCAGGGCCCAGTCCATTGTTCGTGGTACCATTTCTCAGGCACTATTAAAA
AAAAAAAAACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGA
TCTCTGCATCAGCTCCTGCCTCCAGGTTCCTACCCTATTTGAGTTCCTGTCCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTTGCTG
TTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGTTTATAACA
CCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGAAAGTCTCCAGA
CTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCACGTGCACTCCACTGCC
GGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCTGTTTGTTTCCATTTTGACA

Exon 3 (Nuc 8089)

TGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATGATTTTTTTTCCAGGTATAACAGC
CTCACTAGATTGACAGAAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATGCTGCACA
GCAATGGCATTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCAGTCCTTGCAG

Nuc 8218

GTGAGATAGGTAGAGGGTGATGGAGGCTGAGAAGAGAGGTGCAACTGTGGGTTATACCC
AAAAGCTGCTGATTCCCGTGGGAGACATTCTATAAGCATTCTATAAACTAGAGGCAGATA
TCAAGGAAGGATTTCAATTGTAATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATG
CTGGGGAAAATGCTTACACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAA
AAAGCACAGCATGGCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTCAGAAT
CCCAGAGACTCATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAA
ACTTGACTCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACT
CATGCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAGG
AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAAGGAGGGAGGG
AAGGAAGGAAAGGGAAAGGAAAAAAGAGATGGGGAGGGAGGGAAGGAAAGGAAAGGG
GGAGAAAGAAGAGAAGAAAGGAAAATAAATAAATTTTCAGGGATTATTACACCTTTAAA
TTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
AAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAAATTTTGAGTTTTCTCTAC
TTTTTATCCTCTACCATCTTCAAACTGAAACTACAATTATTCCCACAAATGAAGAAAATGC
TGTAAGAGTTTTCACACACCGAAGTGGGAAACTTAAGGATTAGACAAGTCTAACAATGAG
AATGGGGAGAACAAAAAGAGACTGCACAGGGAGCCCTTTCTCTGCTTATAATCTTGACAC
TTGAGAAGCTAATTGACGCTGCATGACTACTCAACTCTTTAAGCAAACAATGCTGTTGTTC
ATGAAAAGCACAATAAAGTACATATGTCCCATAATATTCATCAAAATTTGCATGCAGCAC
ATAATAGCAATCAAAGCAATAACACCCACTGTTCACAGAGACTTTAAACATGAAACTGGA
ACTATGTCTAGTGTTTTGACTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTG
ATTTAGGTCATCAGACAGCATTTGGAACATGTATCTTCAGGAGGAATCATTCATGTATCCT
GCATGAAATTCTCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATT
GGGTTTGAGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG
CTTTTATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTCATGTTCTTTTCTT
TTTTAAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGAAAATAATC
AGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTTTTTGCTTCTAATT
CAGAATCAATTAAATTCATATTACTATAAAAGACAGCACGCCAGATGTGTGCCAGCTGAG
GAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACAGAAAGGAGTGAAGGGTTGA

TGTGCGCTGCAACATCTTGAAAACATTCGGCTACATGATGGAAGCCAGGCACAAAAAGCC
ACATATTGCATGGTTATGTTTATATGAAATGTTTAAAATACATGGATTCTTAGCAAACAGA
GTAAGATGTTACTTAGGGTCAGGAAAAGATTAAAAAAAAAAAAAACTATTGATGTGGAATG
ATCTTAATTTGGGGAAAAGACAATTTCCTAAGACGAAATAGTTGAGGTAGATATAGTTAT
ATCCCTGTGGATATTGTAATAAACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGG
GCAGGAGGAAGTGAAATGAGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTG
GGTGGAATCTAGATATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATC
AAAACCAACAGGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAA
TGTGTGTGGAGACACCATAATAAAACTCCTTTTTTGTTTGCTAACTAAAACCACTAAAATC
TAAAAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGGG
GAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTACGTTT
TTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCATCATCACTA
ATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAGTGAGGCAATCTT
GGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAGGACAGTGAGAGGGAA
ATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCATAGTCCACTACATTACTTTG
TATTCTAATATTAAGAAAATAATAAACCCATTTCTGTGCACTTATCACCCAGGCTCAACAG
TTATCTTGGCCACAGATCCTGTCTCACTGCATCCTGTCCACCTGAGTCCACTTAGCGTTCTG
AATCCAATCCAGGGCATGATGCTTACTCCTACACAGAACTAAAGATTAAAGAGAGTTTAA
AAGTAACCATGACATCTCTCTGTTCCTTTAGCGATAAGTTCTTAATATTTATGGCTGCTTGT
GTATGTTCTAATTTCTCTAATATTGTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGT
TGGAGTTAAGGTCCCATAGGATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCT

Exon 4 ? (Nuc 11286)

TTGTGAAAGTTCCTTTGAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATG
AGCTATAACAAAGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGG
TTGCACCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTTGC
TCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACATTTGTCTCTTT
GAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTTGTATGATAACTTCATTG
ACCTCCCTCCCAAAAGAAATGGTCTCCTCTATGCCAAACCTAGAAAGCCTTTACTTGCATGGAA
ACCCATGGACCTGTGACTGCCATTTAAAGTGGTTGTCCGAGTGGATGCAGGGAAACCCAGGTA

Nuc 11680

ACTATCTTGTTTGTTTGTTTCTTTTTTTATARKACGTATTTTCCTCAATTTCATTTAGAATGA
TATCCCAAAAGTCCCCCATAACCTCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGG
CCCTGGCATTCCCCTGTACTGGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCC
AGTGATGGCCAACTAGGCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGG
TACTGGTTAGTTCATAATGTTGTTGCACCTACAGGGTTGAA

Nuc 11967

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

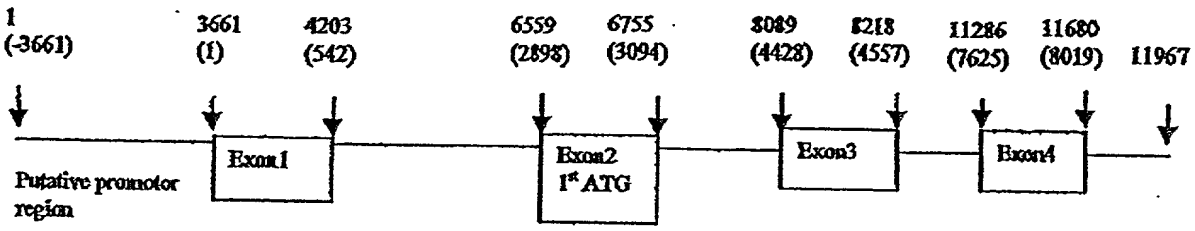
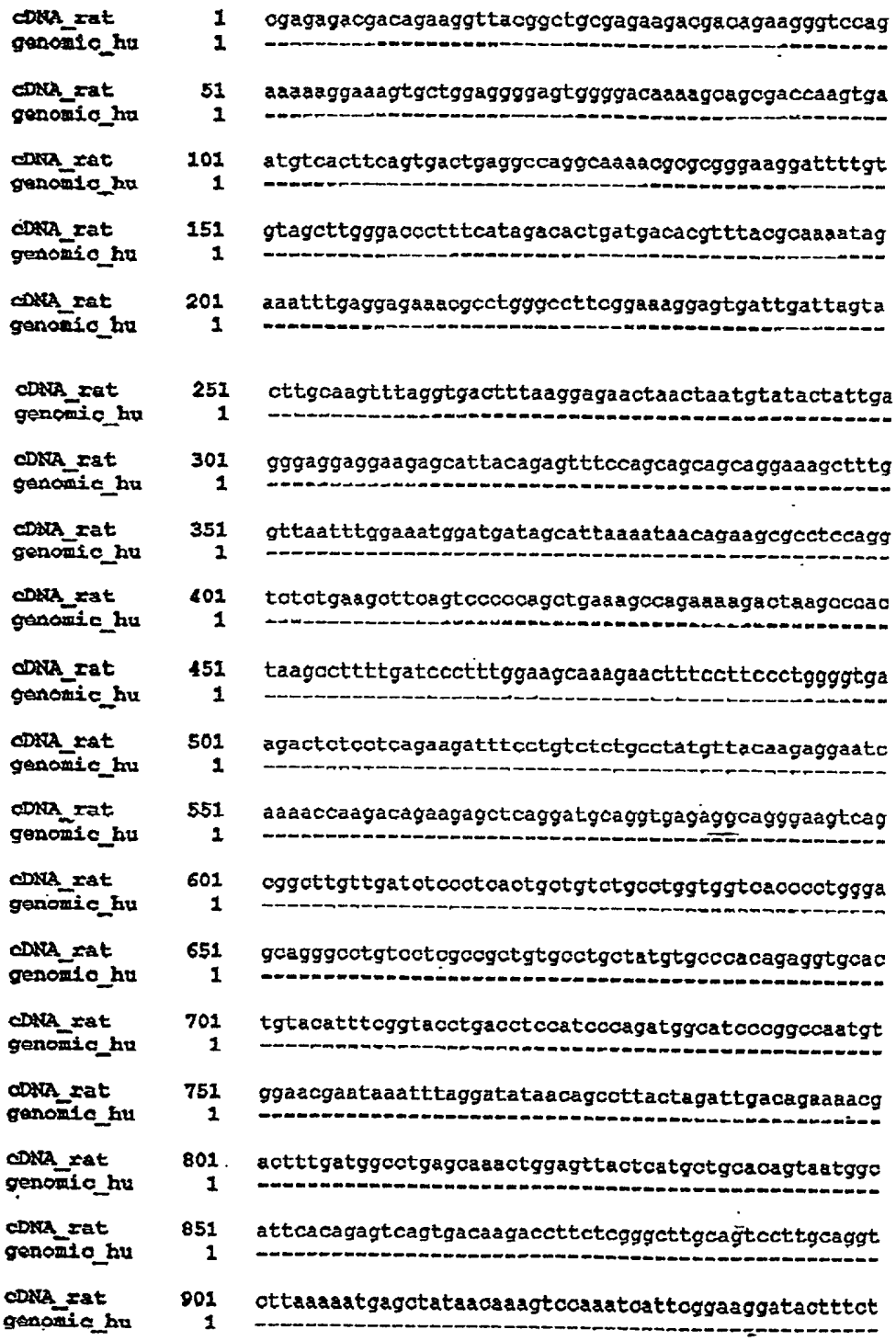


Figure 10


cDNA_rat 1 cgagagacgacagaaggttacggctgcgagaagacgacagaagggtccag
genomic_hu 1 ----------
cDNA_rat 51 aaaaaggaaagtgctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu 1 ----------
cDNA_rat 101 atgtcacttcagtgactgaggccaggcaaaacgcgcgggaaggattttgt
genomic_hu 1 ----------
cDNA_rat 151 gtagcttgggacoctttcatagacactgatgacacgtttacgcaaaatag
genomic_hu 1 ----------
cDNA_rat 201 aaatttgaggagaaacgcctgggccttcggaaaggagtgattgattagta
genomic_hu 1 ----------
cDNA_rat 251 cttgcaagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu 1 ----------
cDNA_rat 301 gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu 1 ----------
cDNA_rat 351 gttaatttggaaatggatgatagcattaaaataacagaagcgcctccagg
genomic_hu 1 ----------
cDNA_rat 401 tctctgaagcttcagtcccccagctgaaagccagaaaagactaagcccac
genomic_hu 1 ----------
cDNA_rat 451 taagcctttgatccctttggaagcaaagaactttccttccctggggtga
genomic_hu 1 ----------
cDNA_rat 501 agactctcctcagaagatttcctgtctctgcctatgttacaagaggaatc
genomic_hu 1 ----------
cDNA_rat 551 aaaaccaagacagaagagctcaggatgcaggtgagaggcagggaagtcag
genomic_hu 1 ----------
cDNA_rat 601 cggcttgttgatctccctcactgctgtctgcctggtggtcacccctggga
genomic_hu 1 ----------
cDNA_rat 651 gcagggcctgtcctcgccgctgtgcctgctatgtgcccacagaggtgcac
genomic_hu 1 ----------
cDNA_rat 701 tgtacatttcggtacctgacctccatcccagatggcatcccggccaatgt
genomic_hu 1 ----------
cDNA_rat 751 ggaacgaataaatttaggatataacagccttactagattgacagaaaacg
genomic_hu 1 ----------
cDNA_rat 801 actttgatggcctgagcaaactggagttactcatgctgcacagtaatggc
genomic_hu 1 ----------
cDNA_rat 851 attcacagagtcagtgacaagaccttctcgggcttgcagtccttgcaggt
genomic_hu 1 ----------
cDNA_rat 901 cttaaaaatgagctataacaaagtccaaatcattcggaaggatactttct
genomic_hu 1 ----------

cDNA_rat	951	acggactcgggagcttggtccggttgcacctggatcacaacaacattgaa
genomic_hu	1	----------
cDNA_rat	1001	ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca
genomic_hu	1	----------
cDNA_rat	1051	tttagaaggaaaccggctcacaaagctccatccagacacatttgtctcat
genomic_hu	1	----------
cDNA_rat	1101	taagctatctccagatatttaaaacctctttcattaagtacctgttcttg
genomic_hu	1	----------
cDNA_rat	1151	tctgataacttcctgacctccctcccaaaagaaatggtctcctacatgcc
genomic_hu	1	----------
cDNA_rat	1201	aaacctagaaagcctgtatttgcatggaaacccatggacctgtgactgcc
genomic_hu	1	----------
cDNA_rat	1251	atttaaagtggttgtctgagtggatgcagggaaacccagatataataaaa
genomic_hu	1	----------
cDNA_rat	1301	tgcaagaaagacagaagctcttccagtcctcagcaatgtccccctttgcat
genomic_hu	1	----------
cDNA_rat	1351	gaaccccaggatctctaaaggcagacccctttgctatggtaccatctggag
genomic_hu	1	----------
cDNA_rat	1401	ctttcctatgtacaaagccaaccattgatccatcactgaagtcaaagagc
genomic_hu	1	----------
cDNA_rat	1451	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt
genomic_hu	1	----------
cDNA_rat	1501	catagaaccctttggctccttgtctttgaacatgacananntntctggaa
genomic_hu	1	----------
cDNA_rat	1551	ataaggccgacatggtctgtagtatccaaaagccatcaaggacatcacca
genomic_hu	1	----------
cDNA_rat	1601	actgcattcactgaagaaaatgactacatcatgctaaatgcgtcattttc
genomic_hu	1	----------
cDNA_rat	1651	cacaaatcttgtgtgcagtgtagattataatcacatccagccagtgtggc
genomic_hu	1	----------
cDNA_rat	1701	aacttctggctttatacagtgactctcctctgatactagaaaggaagccc
genomic_hu	1	----------
cDNA_rat	1751	cagcttaccgagactccttcactgtcttctagatataaacaggtggctct
genomic_hu	1	----------
cDNA_rat	1801	taggcctgaagacatttttaccagcatagaggctgatgtcagagcagacc
genomic_hu	1	----------
cDNA_rat	1851	ctttttggttccaacaagaaaaaattgtcttgcagctgaacagaactgcc
genomic_hu	1	----------
cDNA_rat	1901	accacacttagcacattacagatccagttttccactgatgctcaaatcgc
genomic_hu	1	----------
cDNA_rat	1951	tttaccaagggcggagatgagagcggagagactcaaatggaccatgatcc
genomic_hu	1	----------

cDNA_rat	2001	tgatgatgaacaatcccaaactggaacgcactgtcctggttggcggcact
genomic_hu	1	----------
cDNA_rat	2051	attgccctgagctgtccaggcaaaggcgacccttcacctcacttggaatg
genomic_hu	1	----------
cDNA_rat	2101	gcttctagctgatgggagtaaagtgagagccccttacgttagcgaggatg
genomic_hu	1	----------
cDNA_rat	2151	ggcgaatcctaatagacaaaaatgggaagttggaactgcagatggctgac
genomic_hu	1	----------
cDNA_rat	2201	agctttgatgcaggtctttaccactgcataagcaccaatgatgcagatgc
genomic_hu	1	----------
cDNA_rat	2251	ggatgttctcacatacaggataactgtggtagagccctatggagaaagca
genomic_hu	1	----------
cDNA_rat	2301	cacatgacagtggagtccagcacacagtggttacgggtgagacgctcgac
genomic_hu	1	----------
cDNA_rat	2351	cttccatgcctttccacgggtgttccagatgcttctattagctggattct
genomic_hu	1	----------
cDNA_rat	2401	tccagggaacactgtgttctctcagccatcaagagacaggcaaattctta
genomic_hu	1	----------
cDNA_rat	2451	acaatgggaccttaagaatattacaggttacgccaaaagatcaaggtcat
genomic_hu	1	----------
cDNA_rat	2501	taccaatgtgtggctgccaaccccatcaggggccgacttttccagttttaa
genomic_hu	1	----------
cDNA_rat	2551	agtttcagttcaaaagaaaggccaaaggatggttgagcatgacagggagg
genomic_hu	1	----------
cDNA_rat	2601	caggtggatctggacttggagaaccccaactccagtgtttcccttaagcag
genomic_hu	1	----------
cDNA_rat	2651	ccagcatctttgaaactctctgcatcagctttgacagggtcagaggctgg
genomic_hu	1	----------
cDNA_rat	2701	aaaacaagtctccggtgtacataggaagaacaaacatagagacttaatac
genomic_hu	1	----------
cDNA_rat	2751	atcggcggcgtggggattccacgctccggcgattcagggagcataggagg
genomic_hu	1	----------
cDNA_rat	2801	cagctccctctctctgctcggagaattgacccgcaacgctgggcagcact
genomic_hu	1	----------
cDNA_rat	2851	tctagaaaaagccaaaaagaattctgtgccaaaaaagcaagaaaatacca
genomic_hu	1	----------
cDNA_rat	2901	cagtaaagccagtgccactggctgttcccctcgtggaactcactgacgag
genomic_hu	1	----------
cDNA_rat	2951	gaaaaggatgcctctggcatgattcctccagatgaagaattcatggttct
genomic_hu	1	----------

```
cDNA_rat      3001   gaaaactaaggcttctggtgtcccaggaaggtcaccaactgctgactctg
genomic_hu       1   --------------------------------------------------

cDNA_rat      3051   gaccagtaaatcatggttttatgacgagtatagcttctggcacagaagtc
genomic_hu       1   --------------------------------------------------

cDNA_rat      3101   tcaactgtgaatccacaaacactacaatctgagcaccttcctgatttcaa
genomic_hu       1   --------------------------------------------------

cDNA_rat      3151   attatttagtgtaacaaacggtacagctgtgacaaagagtatgaacccat
genomic_hu       1   --------------------------------------------------

cDNA_rat      3201   ccatagcaagcaaaatagaagatacaaccaaccaaaacccaatcattatc
genomic_hu       1   --------------------------------------------------

cDNA_rat      3251   tttccatcagtagctgaaattcgagattctgctcaggcaggaagagcatc
genomic_hu       1   --------------------------------------------------

cDNA_rat      3301   ttcccaaagtgcacacctgtaacagggggaaacatggctacctatggcc
genomic_hu       1   --------------------------------------------------

cDNA_rat      3351   ataccaacacatatagtagctttaccagcaaagccagtacagtcttgcag
genomic_hu       1   --------------------------------------------------

cDNA_rat      3401   ccaataaatccaacagaaagttatggacctcagatacctattacaggagt
genomic_hu       1   --------------------------------------------------

cDNA_rat      3451   cagcagacctagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu       1   --------------------------------------------------

cDNA_rat      3501   gcttctccagtcacccttcaggttcacacaccactgcctcgtctttattt
genomic_hu       1   --------------------------------------------------

cDNA_rat      3551   cacattcctagaaacaacaatacaggtaacttccccttgtccaggcactt
genomic_hu       1   --------------------------------------------------

cDNA_rat      3601   gggaagagagaggacaatttggagcagagggagagttaaaaacccacata
genomic_hu       1   --------------------------------------------------

cDNA_rat      3651   gaaccccagttctccgacggcatagacacaggactgtgaggccagcaatc
genomic_hu       1   --------------------------------------------------

cDNA_rat      3701   aagggacctgctaacaaaaatgtgagccaagttccagccacagagtaccc
genomic_hu       1   --------------------------------------------------

cDNA_rat      3751   tgggatgtgccacacatgtccttccgcagaggggctcacagtggctactg
genomic_hu       1   --------------------------------------------------

cDNA_rat      3801   cagcactgtcagttccaagttcatcccacagtgccctccccaaaactaat
genomic_hu       1   --------------------------------------------------

cDNA_rat      3851   aatgttggggtcatagcagaagagtctaccactgtggtcaagaaaccact
genomic_hu       1   --------------------------------------------------

cDNA_rat      3901   gttactatttaaggacaaacaaaatgtagatattgagataataacaacca
genomic_hu       1   --------------------------------------------------

cDNA_rat      3951   ctacaaaatattccggaggggaaagtaaccacgtgattcctacggaagca
genomic_hu       1   --------------------------------------------------

cDNA_rat      4001   agcatgacttctgctccaacatctgtatccctggggaaatctcctgtaga
genomic_hu       1   --------------------------------------------------

cDNA_rat      4051   caatagtggtcacctgagcatgcctgggaccatccaaactgggaaagatt
genomic_hu       1   --------------------------------------------------

cDNA_rat      4101   cagtggaaacaacaccacttcccagcccctcagcacaccctcaataccа
genomic_hu       1   --------------------------------------------------
```

cDNA_rat	4151	acaagcacaaaattctcaaagaggaaaactcccttgcaccagatctttgt
genomic_hu	1	----------
cDNA_rat	4201	aaataaccagaagaaggaggggatgttaaagaatccatatcaattcggtt
genomic_hu	1	----------
cDNA_rat	4251	tacaaaagaacccagccgcaaagcttcccaaaatagctcctcttttaccc
genomic_hu	1	----------
cDNA_rat	4301	acaggtcagagttcccccctcagattctacaactctcttgacaagtccgcc
genomic_hu	1	----------
cDNA_rat	4351	accagctctgtctacaacaatggctgccactcagaacaagggcactgaag
genomic_hu	1	----------
cDNA_rat	4401	tagtatcaggtgccagaagtctctcagcagggaagaagcagcccttcacc
genomic_hu	1	----------
cDNA_rat	4451	aactcctctccagtgcttcctagcaccataagcaagagatctaatacatt
genomic_hu	1	----------
cDNA_rat	4501	aaacttcttgtcaacggaaaccccccacagtgacaagtcctactgctactg
genomic_hu	1	----------
cDNA_rat	4551	catctgtcattatgtctgaaacccaacgaacaagatccaaagaagcaaaa
genomic_hu	1	----------
cDNA_rat	4601	gaccaaataaaggggcctcggaagaacagaaacaacgcaaacaccacccc
genomic_hu	1	----------
cDNA_rat	4651	caggcaggtttctggctatagtgcatactcagctctaacaacagctgata
genomic_hu	1	----------
cDNA_rat	4701	ccccttggctttcagtcattccccacgacaagatgatggtggaaatgta
genomic_hu	1	----------
cDNA_rat	4751	agtgcagttgcttatcactcaacaacctctcttctggccataactgaact
genomic_hu	1	----------
cDNA_rat	4801	gtttgagaagtacacccagactttgggaaatacaacagctttggaaacaa
genomic_hu	1	----------
cDNA_rat	4851	cgttgttgagcaaatcacaggagagtaccacagtgaaaagagcctcagac
genomic_hu	1	----------
cDNA_rat	4901	acaccaccaccactcctcagcagtggggcgcccccagtgccoactccttc
genomic_hu	1	----------
cDNA_rat	4951	cccacctcctttactaagggtgtggttacagacagcaaagtcacatcag
genomic_hu	1	----------
cDNA_rat	5001	ctttccagatgacgtcaaatagagtggtcaccatatatgaatcttcaagg
genomic_hu	1	----------
cDNA_rat	5051	cacaatacagatctgcagcaaccctcagcagaggctagccccaatcctga
genomic_hu	1	----------
cDNA_rat	5101	gatcataactggaaccactgactctcoctctaatctgtttccatccactt
genomic_hu	1	----------
cDNA_rat	5151	ctgtgccagcactaagggtagataaaccacagaattctaaatggaagccc
genomic_hu	1	----------
cDNA_rat	5201	tctccctggccagaacacaaatatcagctcaagtcatactccgaaaccat
genomic_hu	1	----------

```
cDNA_rat      5251   tgagaagggcaaaaggccagcagtaagcatgtccccccacctcagccttc
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5301   cagaggccagcactcatgcctcacactggaatacacagaagcatgcagaa
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5351   aagagtgtttttgataagaaacctggtcaaaaccaacttccaaacatct
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5401   gccttacgtctctctacctaagactctattgaaaaagccaagaataattg
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5451   gaggaaaggctgcaagctttacagttccagctaattcagacgttttttctt
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5501   ccttgtgaggctgttggagacccactgcccatcatccactggaccagagt
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5551   ttcatcaggatttgaaatatcccaagggacacagaaaagccggttccacg
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5601   tgcttcccaatggcaccttgtccatccagagggtcagtattcaggaccgt
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5651   ggacagtacctgtgctctgcatttaatccactgggcgtagaccattttca
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5701   tgtctctttgtctgtggtttttacccggcaaggattttggacagacatg
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5751   tcaaggagatcacagttcactttggaagtactgtggaactaaagtgcaga
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5801   gtggagggtatgccgaggcctacggtttcctggatacttgcaaaccaaac
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5851   ggtggtctcagaaacggccaagggaagcagaaaggtctgggtaacacctg
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5901   atggaacattgatcatctataatctgagtctttatgatcgtggtttttac
genomic_hu       1   ---------------------------------------------------

cDNA_rat      5951   aagtgtgtggccagcaacccatctggccaggattcactgttggttaagat
genomic_hu       1   ---------------------------------------------------

cDNA_rat      6001   acaagtcatcacagctccccctgtcattatagagcaaaagaggcaagcca
genomic_hu       1   ---------------------------------------------------

cDNA_rat      6051   tcgttggggttttaggtggaagtttgaaactgccctgcactgcaaaagga
genomic_hu       1   ---------------------------------------------------

cDNA_rat      6101   actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu       1   ---------------------------------------------------

cDNA_rat      6151   accattgcagttgactcattccagatttttcttgtatccaaatggaactc
genomic_hu       1   ---------------------------------------------------

cDNA_rat      6201   tgtatataagaagcatcgctccttcagtgaggggcacttatgagtgcatt
genomic_hu       1   ---------------------------------------------------

cDNA_rat      6251   gccaccagctcctcaggctcagagagaagggtagtgattcttactgtgga
genomic_hu       1   ---------------------------------------------------
```

:at	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
ic_hu	890	CtaTGGAAGCAGAATCACAGTCCATAAAAATGGAACCTTGGAaattagga
:at	7401	acatccGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGttcggagcgag
.c_hu	940	atgtgaGGCTTTCAGATTCAGCCGACTTTATCTGTGTGGcccgaaatgaa
:at	7451	ggaGGAGAGAGTGTGTTGGTAGTGCAGTTAGAAGTCCTAGAAATGCTGAG
.c_hu	990	ggtGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAG
:at	7501	AAGACCAACATTCAGAAACCCATTCAACGAAAAagtcatcgcccaagctg
.c_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgcccagctgg
:at	7551	gcaagcccgtAGCACTGAACTGCTCTGTGGATGGGAACCCCCCACCTGAA
.c_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAA
:at	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAcacaa
.c_hu	1140	ataatCTGGATTTTACCAAATGGCACACGATTTTCCAATGGACCAcaaag
:at	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAAgcaa
.c_hu	1190	ttatcaGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAacaa
:at	7701	ctcggaacaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGC
.c_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGC
:at	7751	TAcatcgagaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
.c_hu	1290	TAtattgagaaattagtcatattagaaATTGGCCAGAAGCCAGTTATTCT
:at	7801	GACATACGAACCAGGGAtggtgaagagcgTCAGTGGGGAACCGTTATCAC
.c_hu	1340	TACCTATGCACCAGGGAcagtaaaaggcaTCAGTGGAGAATCTCTATCAC
:at	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
.c_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
:at	7901	CCGGGTGGccATGTAATCGACAGGCCTCAAGTGGATGGAAAATACATAcT
.c_hu	1440	CCAAGTGGttATGTAGTAGACAGGCCTCAAATTAATGGGAAATACATAtT
:at	7951	GCATGAAAATGGCACGCTGGTCATCAAAGCAACAACAGCtcacgaccaAG
ic_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
:at	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTTGGCCAggcagttattagc
.c_hu	1540	GAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCAtacactgattact
:at	8051	gtgtCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAAActacctACC
.c_hu	1590	gttcCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAAtcgtccACC
:at	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
.c_hu	1640	CAGGAGTATTGTCACCAGGACAGGGGCAGCCtttCAGCTCCACTGTGTGG
:at	8151	CCTTGGGAATCCCCAAGCCAaAAGTCACCTGGGAGACGCCAAGACACTCC
.c_hu	1690	CCTTGGGAGTTCCCAAGCCAgAAATCACATGGGAGATGCCTGACCACTCC
:at	8201	CTGCTCTCAAaagcaacagcaagaaaacccCATAGAAGTGAGATGCTTCA
.c_hu	1740	CTTCTCTCAAcggcaagtaaagagaggacaCATGGAAGTGAGCAGCTTCA
:at	8251	CCCACAAGGTACGCTgGTCATTCAGAATCTCCAAACCTCGGATTCCGGag
.c_hu	1790	CTTACAAGGTACCCTaGTCATTCAGAATCCCCAAACCTCCGATTCTGGga
:at	8301	tcTATAAGTGCAGAGCTCAGAACCTACTTGGgacTGATTACGCAACAACT
.c_hu	1840	taTACAAATGCACAGCAAAGAACCCACTTGGtagTGATTATGCAGCAACG
:at	8351	TACATCCAGGTACTCTGACAGGAAgggggagactaaaattcaacagaagt
.c_hu	1890	TATATTCAAGTAATCTGACATGAAataataaagtcaacaacatctgggca

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~2317	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

Figure 12

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCCTGTCCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACTGGAGTTACTCATGCTTCACAGCAATGGC
ATTCACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCACTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAATC
TCTTTCATTAAGTTCCTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGCCAGATGTAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCTCAGCAGTGT
CCACTTTGCATGAACCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAG
TGTGCCAAGCCAACCATTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCT
GCTTTCATCTCTCCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAGTCT
GGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCCATTGCATTCACTGAA
GAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTTGGTGTGCAACATAGATTACGGTCAC
ATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTG
CTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTTACC
AACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCCTTGCAGCTGAAC
AGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGA
GCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACAT
ACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCCACCCCACACGTGGAT
TGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCTTATGTCAGTGAGGATGGACGGATCCTAATAGAC
AAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTGACACAGGCGTATATCACTGTATAAGCAGC
AATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTCGAAGCCTATCAG
GAAAATGGGATTCATCACACAGTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACTGGTATC
CCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAA
GTTCTAAACAATGGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTTGATTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAAGGACAAAGGCCC
TTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTTAAGGAG
CCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACA
AGTAAGAGGCACAACTATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGG
GAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTGTTGGAG
AAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCACCCCCAGTGGTCACC
CAACTCCCAAACATACCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATTTATG
GTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGAT
AGTCCTATGACAAACATAAATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCT
GAAGAACCCACAGATTTCAAACTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATAAACCCA
ACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCATCCACTGTCTTTCCACTGCTACTTGGA
GCAACTGAATTTCAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCCAATA
ACAGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGCACCACCAACAAACTATTATTA
GAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCAC
TTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGCTCAT
TCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTCAGACGCTTTGGGAGGCAGAGG
AAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGC
ATTTTCAGGTCAACAACCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAAT
GTGACATGTCTGTCCTGTCTTCCCAGGGAGAGGCTCACCACTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACTCTAGTC
CAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAAACAACACCCACAATAAAATATTTC
AGGACTGAAATTTCCCAAGTGACTCCAACTGGTGCAGTCATGACATATGCTCCAACATCCATACCCATG
GAAAAAACTCACAAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCA
GTGATTACATCGTCACTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGGTTT
TCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGAAT
CAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCCTAAAACATCTCCTGCTTTACCACAG
AGACAAAGTTCCCCTTTCCATTTCACCACACTTTCAACAAGTGTGATGCAAATTCCATCTAATACCTTG
ACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCC
TTCCCACCCCTTAACCCTATGCTTCCTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATCATATCA
ACGCAAACAGCAATACCAGCAACAACTCCTACCTTCCCTGCATCTGTCATCACTTATGAAACCCAAACA
GAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAGCCTCAAAAGAAGAACAGGACTGACCCAAAC
ATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACT

CATTCCCCACCAGAAAACACAACTGGGATTTCAAGCACAATCAGTTTTCATTCAAGAACTCTTAATCTG
ACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCTGAAACA
ACTTTGTCCAGCAAATCACACCAGAGTACCACAACTAGGAAAGCATCATTAGACACTCCCATACCACCA
TTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCAGTTACT
GACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGG
CACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACTCCATCTCCCTGGGCAGAA
TACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGCATGTTG
GCTACTACAGGCCTGTCCGAGGCCACCACTCTTGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAG
AGTGACTTTGATAAGAAACCAGTTCAAGAAGCAACAACTTCCAAACTCCTTCCCTTTGACTCTTTGTCT
AGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCTAACTCA
GATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCCACCATTCATTGGACCAGAGTTTCAGGA
CTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCTGTCCATCCAG
AGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTT
CATGTCACCTTGTCTGTGGTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTT
CATTCCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCTGGATT
CTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGA
ACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCAGGT
GGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAAAGGAGG
CAAGTCATTGTAGGCACTTGGGGTGAAAGTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCC
AGCGTTTACTGGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCAAGTTGTTC
TTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATT
GCTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCC
AGGATAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAACTGCTCA
GCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCGACCAGTGGAGC
TGGATCCACGTCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTAC
TTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAACCT
GCCAAAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGC
AAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAACAATGCA
ATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACTTTATACTTC
AACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGCTATGCCCAGAACACCCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAGTAACAAAACCAACAAGAGA
ATCAAAGCTGGAGACACAGCTGTCCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGG
TTGCTGCCTTCCAATGACATGATTTCCTTCTCCATTGATAGGTACACATTTCATGCCAATGGGTCTTTG
ACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGTGGGGAT
GACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACAAACAGA
ACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCTGAAGGGACACCA
TCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCCTCACAGCCCCATACTATGGAAGCAGAATC
ACAGTCCATAAAAATGGAACCTTGGAAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGT
GTGGCCCGAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGA
CCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGAAAGTCCACAGCATTGAATTGC
TCTGTTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCCAATGGA
CCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAACAACTCGGGAGGAT
GCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCA
CTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTATGTAGTA
GACAGGCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACA
GCTTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCATACACTGATTACTGTTCCA
GTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACAGGG
GCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGGAGATGCCTGAC
CACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTGAGCAGCTTCACTTACAAGGTACC
CTAGTCATTCAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGT
AGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 13

A

Region						
General	Rat	Human	Region Length	% identity	% positives	% gaps
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region					
Rat	Mouse	Length	% identity	% positives	% gaps
1-238	1-238	238	91	92	1

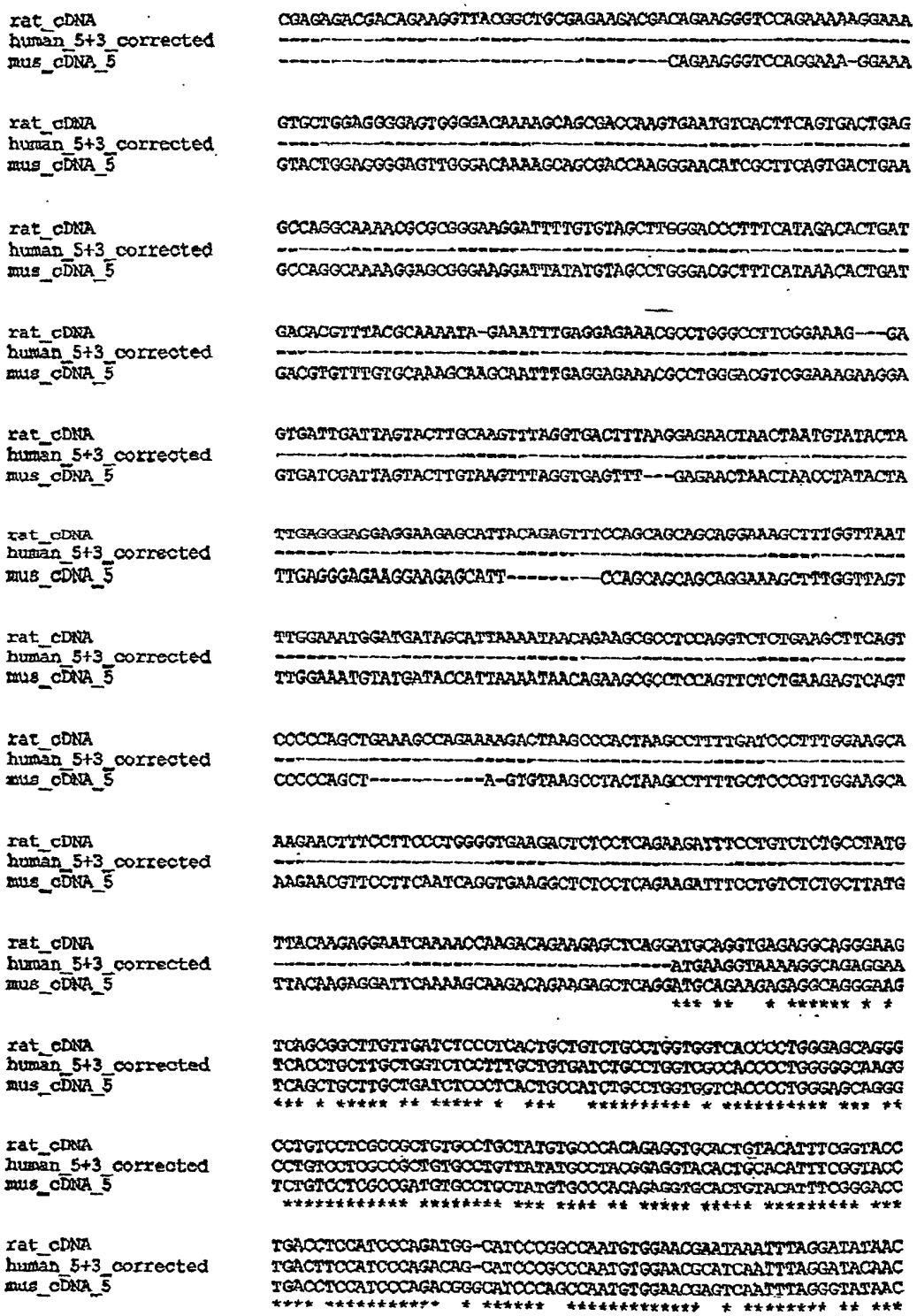
C

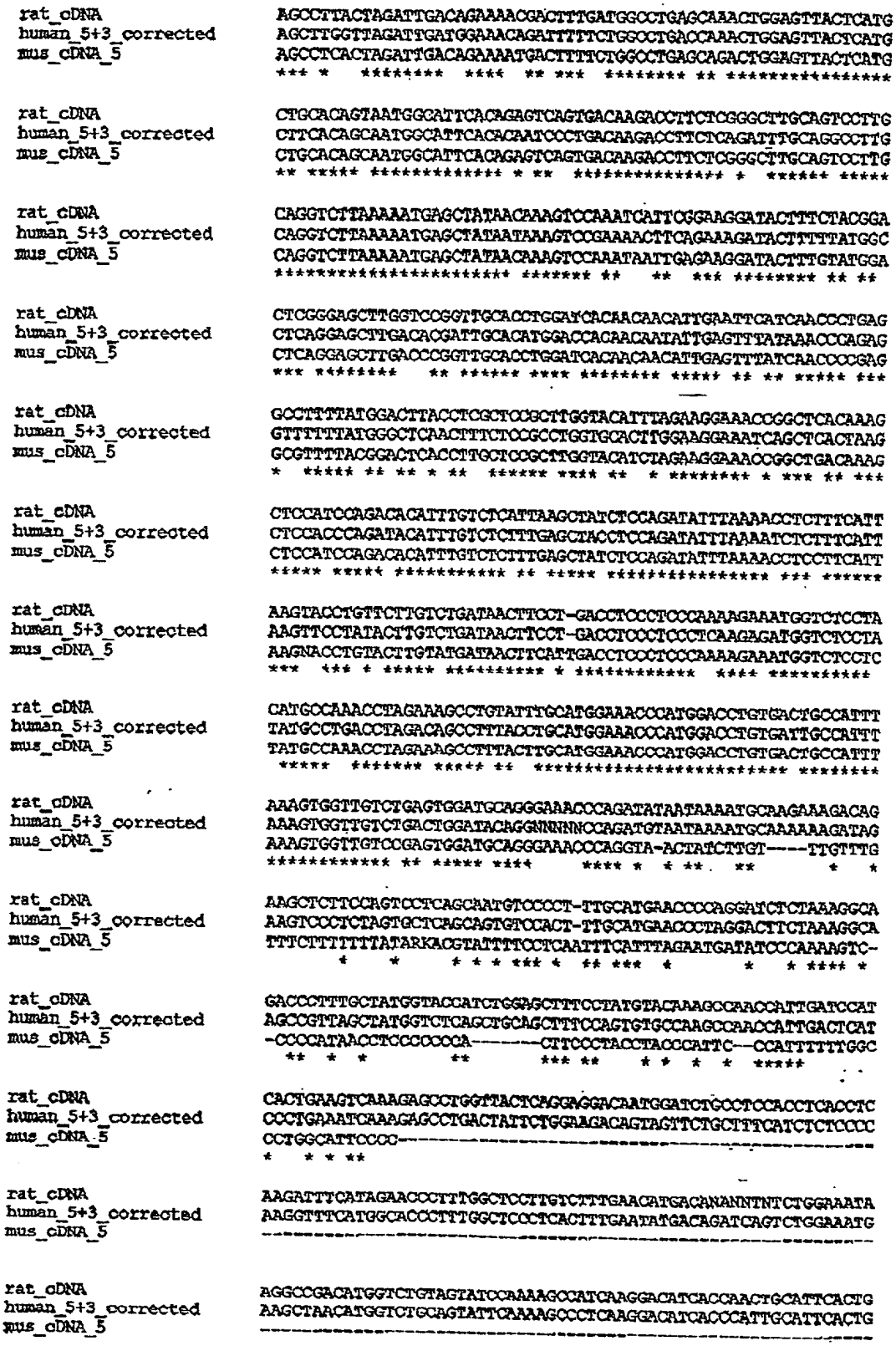
Region				
General	Rat	Human	Region Length	% identity
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

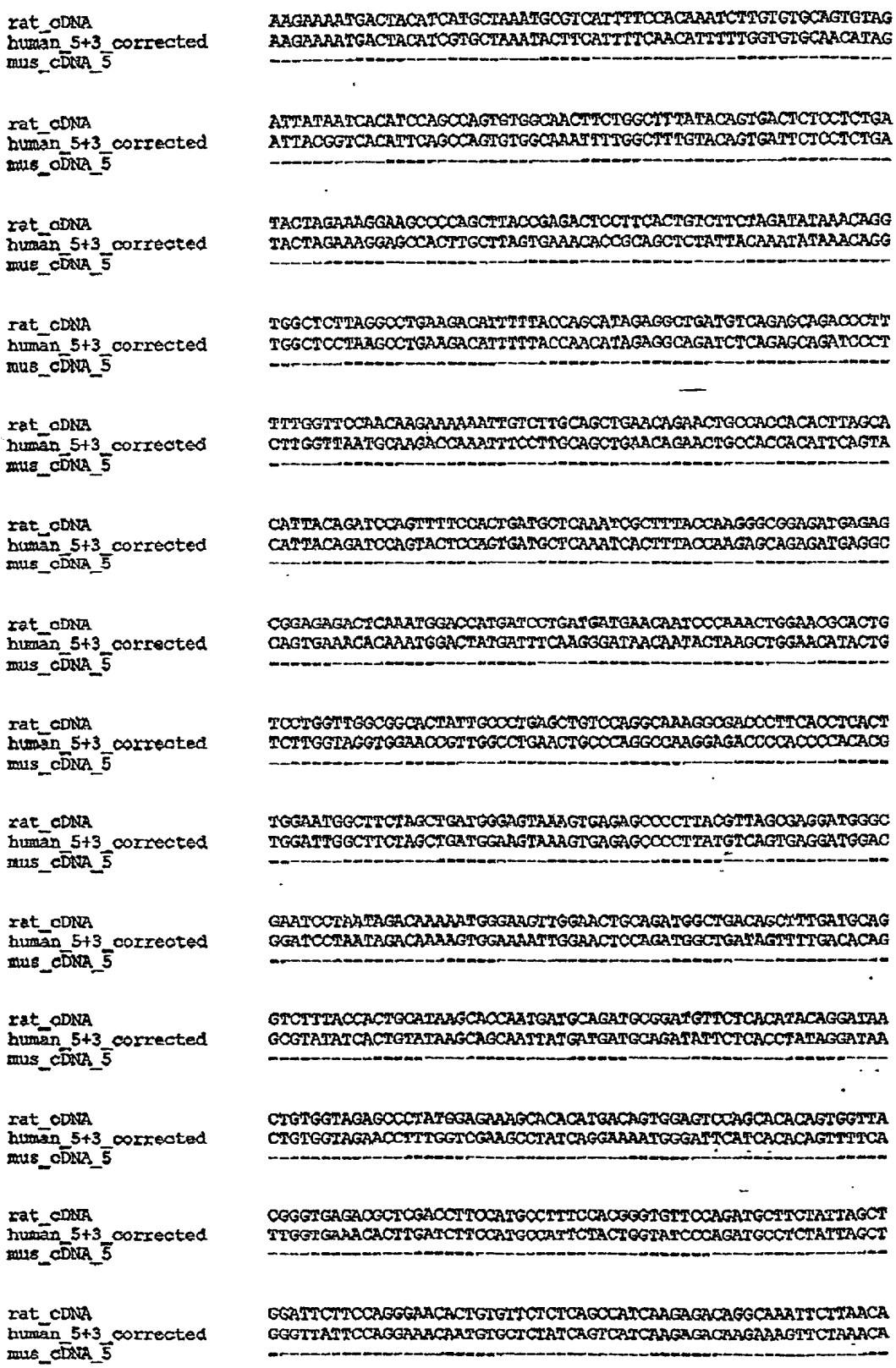
D

Region				
General	Rat	Mouse	Region Length	% identity
1-720	1-718	1-720	720	93

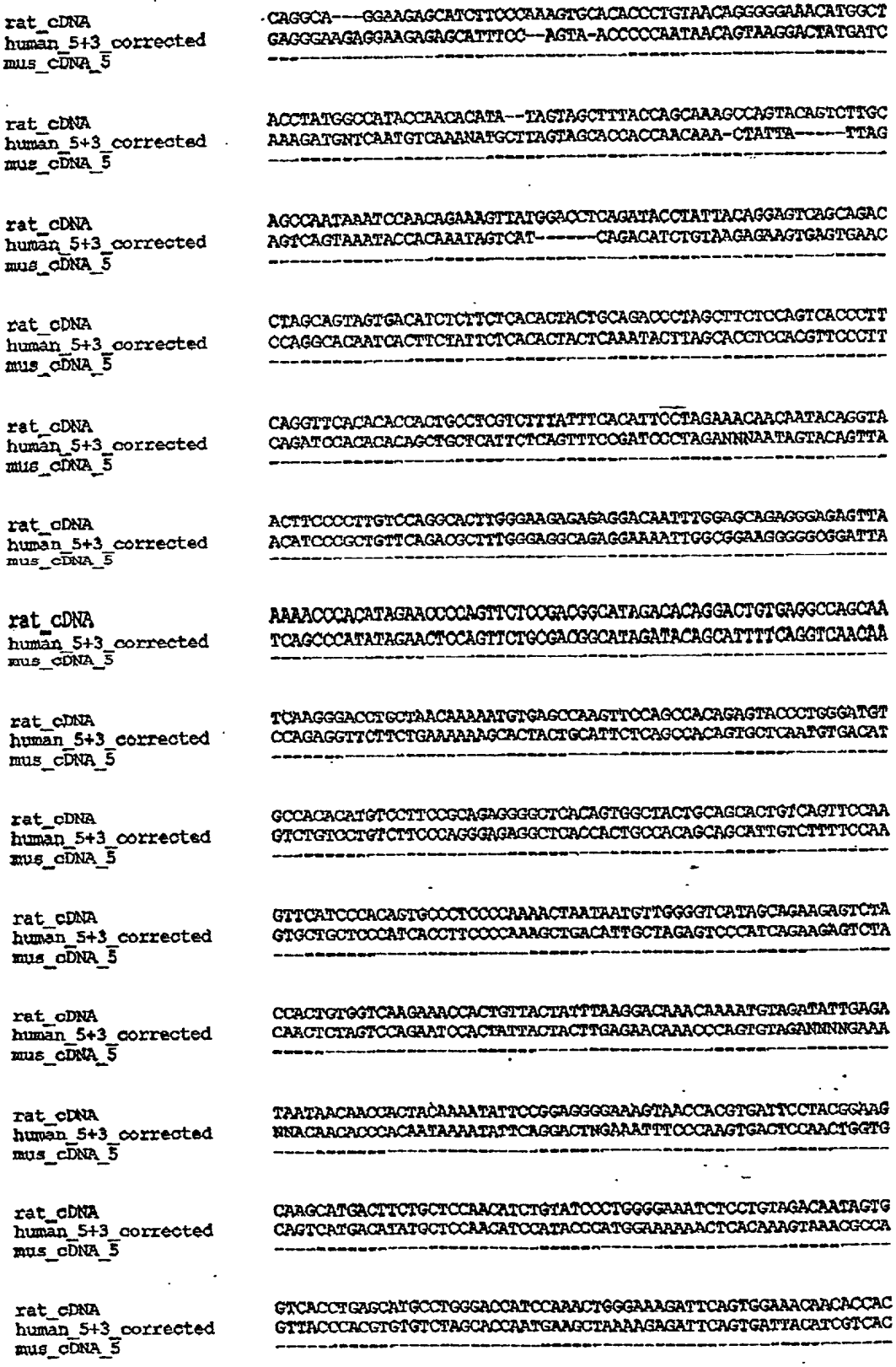
Figure 14


rat_cDNA
human_5+3_corrected
mus_cDNA_5
CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAAGGAAA
----------CAGAAGGGTCCAGGAAA-GGAAA
GTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG
GTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAA
GCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTTTCATAGACACTGAT
GCCAGGCAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGAT
GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCCTGGGCCTTCGGAAAG---GA
GACGTGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA
GTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA
GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTT---GAGAACTAACTAACCTATACTA
TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGCAGCAGGAAAGCTTTGGTTAAT
TTGAGGGAGAAGGAAGAGCATT---------CCAGCAGCAGCAGGAAAGCTTTGGTTAGT
TTGGAAATGGATGATAGCATTAAAATAACAGAAGCGCCTCCAGGTCTCTGAAGCTTCAGT
TTGGAAATGTATGATACCATTAAAATAACAGAAGCGCCTCCAGTTCTCTGAAGAGTCAGT
CCCCCAGCTGAAAGCCAGAAAAGACTAAGCCCACTAAGCCTTTTGATCCCTTTGGAAGCA
CCCCCAGCT----------A-GTGTAAGCCTACTAAGCCTTTGCTCCCGTTGGAAGCA
AAGAACTTTCCTTCCCTGGGGTGAAGACTCTCCTCAGAAGATTTCCTGTCTCTGCCTATG
AAGAACGTTCCTTCAATCAGGTGAAGGCTCTCCTCAGAAGATTTCCTGTCTCTGCTTATG
TTACAAGAGGAATCAAAACCAAGACAGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAG
----------ATGAAGGTAAAAGGCAGAGGAA
TTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGGATGCAGAAGAGAGGCAGGGAAG
TCAGCGGCTTGTTGATCTCCCTCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGG
TCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACCCCTGGGGGCAAGG
TCAGCTGCTTGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGG
CCTGTCCTCGCCGCTGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGTACC
CCTGTCCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCGGTACC
TCTGTCCTCGCCGATGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGGACC
TGACCTCCATCCCAGATGG-CATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAAC
TGACTTCCATCCCAGACAG-CATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAAC
TGACCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTATAAC

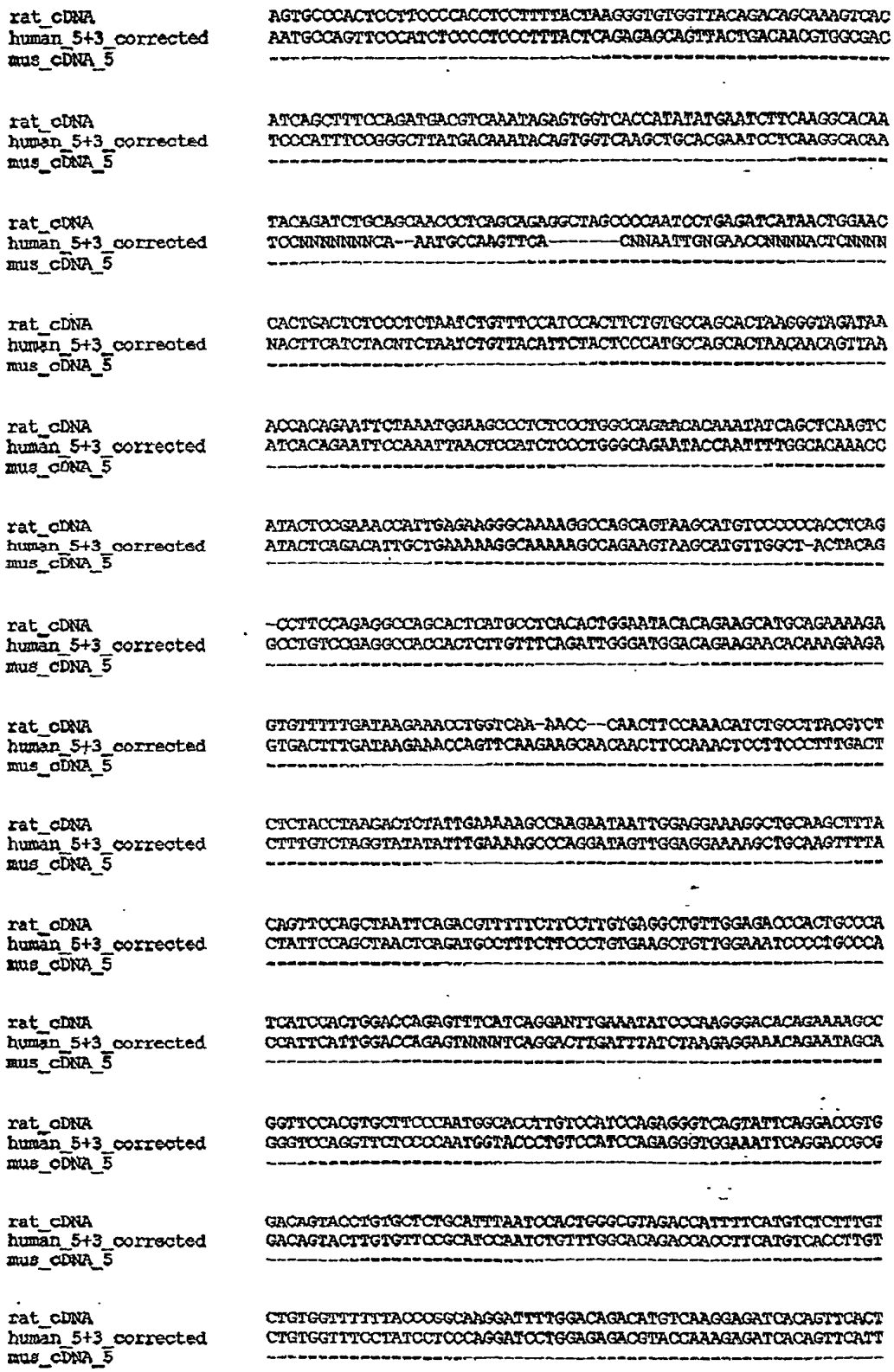

rat_cDNA
human_5+3_corrected
mus_cDNA_5

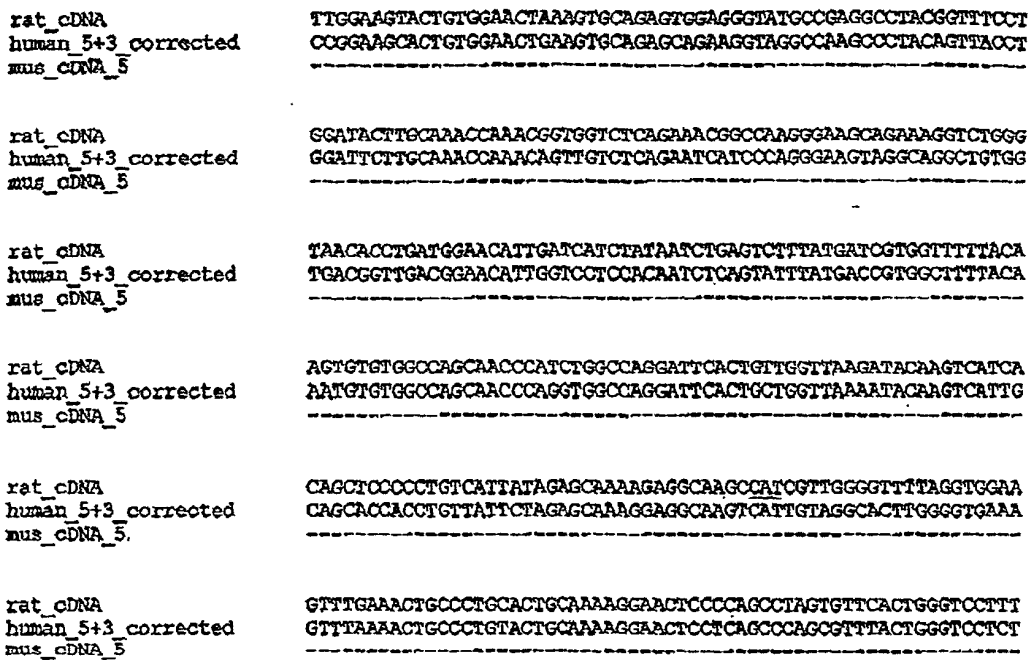

rat_cDNA
human_5+3_corrected
mus_cDNA_5
AAGAAAATGACTACATCATGCTAAATGCGTCATTTTCCACAAATCTTGTGTGCAGTGTAG
AAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTTGGTGTGCAACATAG
ATTATAATCACATCCAGCCAGTGTGGCAACTTCTGGCTTTATACAGTGACTCTCCTCTGA
ATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGA
TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGG
TACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGG
TGGCTCTTAGGCCTGAAGACATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTT
TGGCTCCTAAGCCTGAAGACATTTTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCT
TTTGGTTCCAACAAGAAAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCA
CTTGGTTAATGCAAGACCAAATTTCCTTGCAGCTGAACAGAACTGCCACCACATTCAGTA
CATTACAGATCCAGTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTACCAAGAGCAGAGATGAGGC
CGGAGAGACTCAAATGGACCATGATCCTGATGATGAACAATCCCAAACTGGAACGCACTG
CAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACATACTG
TCCTGGTTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACT
TCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCCACCCCACACG
TGGAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGGC
TGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCTTATGTCAGTGAGGATGGAC
GAATCCTAATAGACAAAAATGGGAAGTTGGAACTGCAGATGGCTGACAGCTTTGATGCAG
GGATCCTAATAGACAAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTGACACAG
GTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATACAGGATAA
GCGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA
CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA
CTGTGGTAGAACCTTTGGTCGAAGCCTATCAGGAAAATGGGATTCATCACACAGTTTTCA
CGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAAACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGCT
GGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAGACAGGCAAATTCTTAACA
GGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAACA

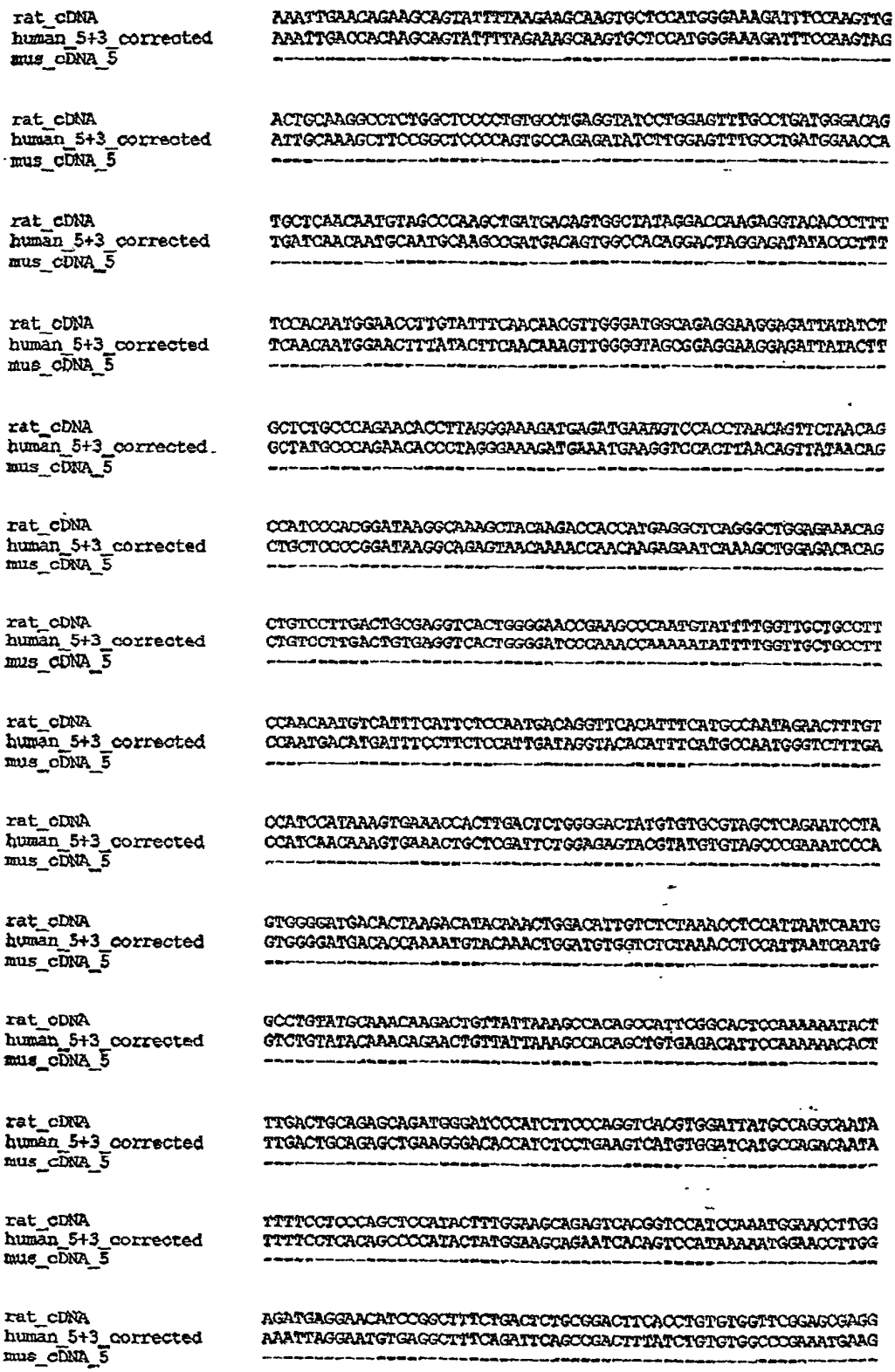
```
rat_cDNA                ATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG
human_5+3_corrected     ATGGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTGG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTTAAAGTTTCAGTTCAAAAGAAAGGCC
human_5+3_corrected     CAGCCAACCCATCAGGGGTTGATTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAAGGAC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AAAGGATGGTTGAGCATGACAGGGAGGCAGGTGGATCTGGACTTGGAGAACCCAACTCCA
human_5+3_corrected     AAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GTGTTTCCCTTAAGCAGCCAGCATCTTTGAAACTCTCTGCATCAGCTTTGACAGGGTCAG
human_5+3_corrected     TTGCTCATCTTAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AGGCTGGAAAACAAGTCTCCGGTGTACATAGGAAGAACAAACATAGAGACTTAATACATC
human_5+3_corrected     AGGTTGGAAAACACACCTCAAGCACAAGTAAGAGGCACAACTATCGGGAATTAACACTCC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                GGCGGCGTGGGGATTCCACGCTCGGCGATTCAGGGAGCATAGGAGGCAGCTCCCTCTCT
human_5+3_corrected     AGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGGAGGCATTTCCCTCCCT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CTGCTCGGAGAATTGACCCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATT
human_5+3_corrected     CTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTGTTGGAGAAAGCTAAAAAGAATG
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CTGTGCCAAAAAAGCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCG
human_5+3_corrected     CTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCACCCCCAGTGGTCACCCAACTCC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TGGAACTCACTGACGAGGAAAAGGATGCCTCTGGCATGATTCCTCCAGATGAAGAATTCA
human_5+3_corrected     CAAACATACCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAATTTA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                TGGTTCTGAAAACTAAGGCTTCTGGTGTCCCAGGAAGGTCACCAACTGCTGACTCTGGAC
human_5+3_corrected     TGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CAGTAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
human_5+3_corrected     CAATATCTGATAGTCCTATGACAAACATAAATTATGGCACAGAACTCTCCGTTGTGAATT
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CACAAACACTACAATCTGAGCACCTTCCTGATTTCAAATTATTTAGTGTAACAAACGGTA
human_5+3_corrected     CACAAATACTACCACCTGAAGAACCCACAGATTTCAAACTGTCTACTGCTATTAAAACTA
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACAACCAACC
human_5+3_corrected     CAGCCATGTCAAAGAATATAAACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATC
mus_cDNA_5              ------------------------------------------------------------

rat_cDNA                AAAACCCAATCATTATCTTTCCATC---------AGTAGCTGAAATTCGAGATTCTG-CT
human_5+3_corrected     AACATTCATCCACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTGACA
mus_cDNA_5              ------------------------------------------------------------
```

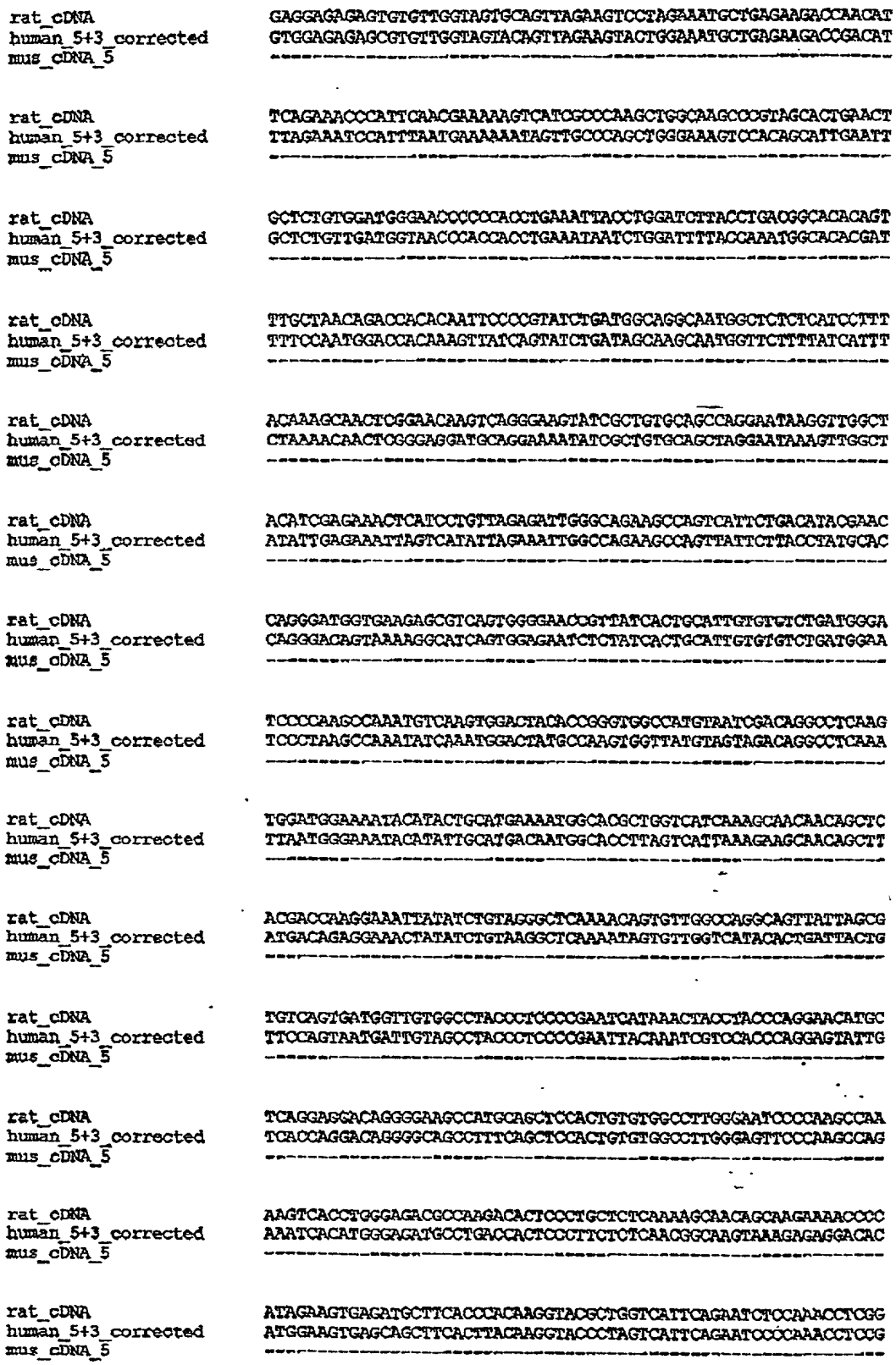

rat_cDNA
human_5+3_corrected
mus_cDNA_5
CAGGCA---GGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGAAACATGGCT
GAGGGAAGAGGAAGAGAGCATTTCC--AGTA-ACCCCCAATAACAGTAAGGACTATGATC
ACCTATGGCCATACCAACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTGC
AAAGATGNTCAATGTCAAANATGCTTAGTAGCACCACCAACAAA-CTATTA-----TTAG
AGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCAGCAGAC
AGTCAGTAAATACCACAAATAGTCAT------CAGACATCTGTAAGAGAAGTGAGTGAAC
CTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCCTAGCTTCTCCAGTCACCCTT
CCAGGCACAATCACTTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTT
CAGGTTCACACACCACTGCCTCGTCTTTATTTCACATTCCTAGAAACAACAATACAGGTA
CAGATCCACACACAGCTGCTCATTCTCAGTTTCCGATCCCTAGANNNAATAGTACAGTTA
ACTTCCCCTTGTCCAGGCACTTGGGAAGAGAGAGGACAATTTGGAGCAGAGGGAGAGTTA
ACATCCCGCTGTTCAGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGCGGATTA
AAAACCCACATAGAACCCCAGTTCTCCGACGGCATAGACACAGGACTGTGAGGCCAGCAA
TCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTTCAGGTCAACAA
TCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCCAGCCACAGAGTACCCTGGGATGT
CCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACAT
GCCACACATGTCCTTCCGCAGAGGGGCTCACAGTGGCTACTGCAGCACTGTCAGTTCCAA
GTCTGTCCTGTCTTCCCAGGGAGAGGCTCACCACTGCCACAGCAGCATTGTCTTTTCCAA
GTTCATCCCACAGTGCCCTCCCCAAAACTAATAATGTTGGGGTCATAGCAGAAGAGTCTA
GTGCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTA
CCACTGTGGTCAAGAAACCACTGTTACTATTTAAGGACAAACAAAATGTAGATATTGAGA
CAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGANNNGAAA
TAATAACAACCACTACAAAATATTCCGGAGGGGAAAGTAACCACGTGATTCCTACGGAAG
NNACAACACCCACAATAAAATATTCAGGACTNGAAATTTCCCAAGTGACTCCAACTGGTG
CAAGCATGACTTCTGCTCCAACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTG
CAGTCATGACATATGCTCCAACATCCATACCATGGAAAAAACTCACAAAGTAAACGCCA
GTCACCTGAGCATGCCTGGGACCATCCAAACTGGGAAAGATTCAGTGGAAACAACACCAC
GTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCAGTGATTACATCGTCAC

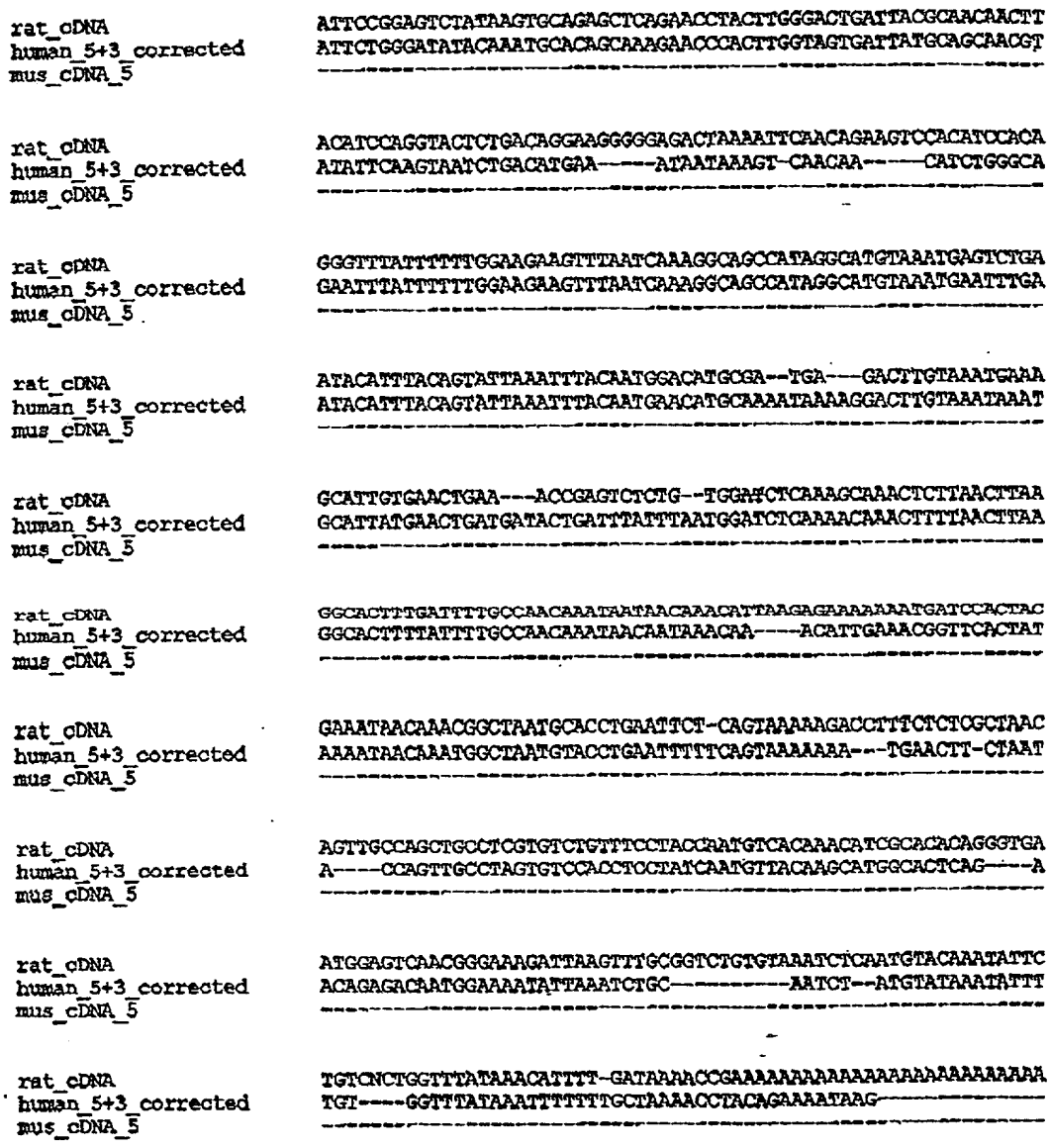
```
rat_cDNA                 TTCCCAGCCCCCTCAGCACACCCTCAATACCAA------CAAGCACAAAATTCTCAAAGA
human_5+3_corrected      TTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGGTTTTCAAGAA
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 GGAAAACTCCCTTGCACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGA
human_5+3_corrected      GGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTAAGGA
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 ATCCATATCAATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAAATAGCTCCTC
human_5+3_corrected      ATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCCTAAAACATCTCCTG
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 TTTTACCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACTCTCTTGACAAGTCCGCCA
human_5+3_corrected      CTTTACC-ACAGAGACAAAGTTCCCCTTTCCATTTCACCACACTTTCAACAAGTGTGATG
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGT
human_5+3_corrected      CAAATTCCATCTAATACCTTGACTACCGCTCACCACACTACGACCAAA--ACACACAA-T
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 GCCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAG-TGCTTCC
human_5+3_corrected      CCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTC-CCACCCCTTAACCCTATGCTTCC
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 TAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCACAGT
human_5+3_corrected      TAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATCATATCAACGCAAACAGCAACCGC
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 GACAAGTCCTACTGCTACTGCATCTGTCATTATGTCTGAAACCCAACGAACAAGATCCAA
human_5+3_corrected      AACAACTCCTACCTTCCCTGCATCTGTCATCACTTATGAAACCCAAACAGAGAGATCTAG
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 AGAAGCAAAAGACCAAATAAAGGGGCCTCG---GAAGAACAGAAACAACGCAAACACCAC
human_5+3_corrected      AGCACAAACAATACAAAGAGAAGGACCTCAAAAGAAGAACAGGACTGACCCAAACATCTC
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 CCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGCTGATACCCCCTT
human_5+3_corrected      TCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACNACCTCCTNNGCTCTNNN
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 GGCTTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAGTGCAGTTGCTTATCA
human_5+3_corrected      NGCATTCACTCATTCCCCACCAGAAAACACAACTGGGATTTCAAGCACAATCAGTTTTCA
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 CTCAACAACCTCTCTTCTGGCCA---TAACTGAACTGTTTGAG-AAGTAC--ACCCAGAC
human_5+3_corrected      TTCAAGAACTCTTAATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGAC
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 TTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAATCACAGGAGAGTACCAC
human_5+3_corrected      TTTGAAGAGCACAATTGCTTCTGAAACAACTTTGTCCAGCAAATCACACCAGAGTACCAC
mus_cDNA_5               ------------------------------------------------------------

rat_cDNA                 AGTGAAAAGAGCCTCA---GACAC---ACCACCACCACTCCTCAGCAGTGGGGCGCCCCC
human_5+3_corrected      AACTAGGAAAGCATCATTAGACACTCAACCACCACCATTCTTGAGCAGCAGTGCTACTCT
mus_cDNA_5               ------------------------------------------------------------
```


rat_cDNA
human_5+3_corrected
mus_cDNA_5
AGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTTACAGACAGCAAAGTCAC
AATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCAGTTACTGACAACGTGGCGAC
ATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA
TCCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAA
TACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATAACTGGAAC
TCCNNNNNNNCA--AATGCCAAGTTCA-------CNNAATTGNGAACCNNNNACTCNNNN
CACTGACTCTCCCTCTAATCTGTTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA
NACTTCATCTACNTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAACAACAGTTAA
ACCACAGAATTCTAAATGGAAGCCCTCTCCCTGGCCAGAACACAAATATCAGCTCAAGTC
ATCACAGAATTCCAAATTAACTCCATCTCCCTGGGCAGAATACCAATTTTGGCACAAACC
ATACTCCGAAACCATTGAGAAGGCAAAAGGCCAGCAGTAAGCATGTCCCCCCACCTCAG
ATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGCATGTTGGCT-ACTACAG
-CCTTCCAGAGGCCAGCACTCATGCCTCACACTGGAATACACAGAAGCATGCAGAAAAGA
GCCTGTCCGAGGCCACCACTCTTGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGA
GTGTTTTTGATAAGAAACCTGGTCAA-AACC--CAACTTCCAAACATCTGCCTTACGTCT
GTGACTTTGATAAGAAACCAGTTCAAGAAGCAACAACTTCCAAACTCCTTCCCTTTGACT
CTCTACCTAAGACTCTATTGAAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTA
CTTTGTCTAGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTA
CAGTTCCAGCTAATTCAGACGTTTTCTTCCTTGTGAGGCTGTTGGAGACCCACTGCCCA
CTATTCCAGCTAACTCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCCA
TCATCCACTGGACCAGAGTTTCATCAGGANTTGAAATATCCCRAGGGACACAGAAAAGCC
CCATTCATTGGACCAGAGTNNNNTCAGGACTTGATTTATCTAAGAGGAAACAGAATAGCA
GGTTCCACGTGCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTG
GGGTCCAGGTTCTCCCCAATGGTACCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCG
GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGTCTCTTTGT
GACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGT
CTGTGGTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACT
CTGTGGTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCATT


rat_cDNA
human_5+3_corrected
mus_cDNA_5
TTGGAAGTACTGTGGAACTAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTCCT
CCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACCT
GGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGGTCTGGG
GGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGG
TAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA
TGACGGTTGACGGAACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACA
AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGATACAAGTCATCA
AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTG
CAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTGGGGTTTTAGGTGGAA
CAGCACCACCTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGCACTTGGGGTGAAA
GTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTAGTGTTCACTGGGTCCTTT
GTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGCGTTTACTGGGTCCTCT
ATGATGGGACTGAACTAAAACCATTGCAGTTGACTCATTCCAGATTTTTCTTGTATCCAA
CTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCAAGTTGTTCTTATTTTCAA
ATGGAACTCTGTATATAAGAAGCATCGCTCCTTCAGTGAGGGGCACTTATGAGTGCATTG
ATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAATGCATTG
CCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA
CTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA
CAATCCCCAGGATAGAAACTGCCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT
CCAGCCCCAGGATAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTGGGGACAAAT
TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT
TACTACTGAACTGCTCAGCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCAT
CCAAGGCTGTCATCGACCAGTGGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATG
CCAAGGCTGTGGTCGACCAGTGG----------GCAGCTGGATCCACGTCTACCCTAATG
GATCCTTGGTGGTTGGGTCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAA
GATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTCTACTTGTGTGTGGCAA
GAAACAAAATGGGAGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCA
GAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTGAAACCTGCCA

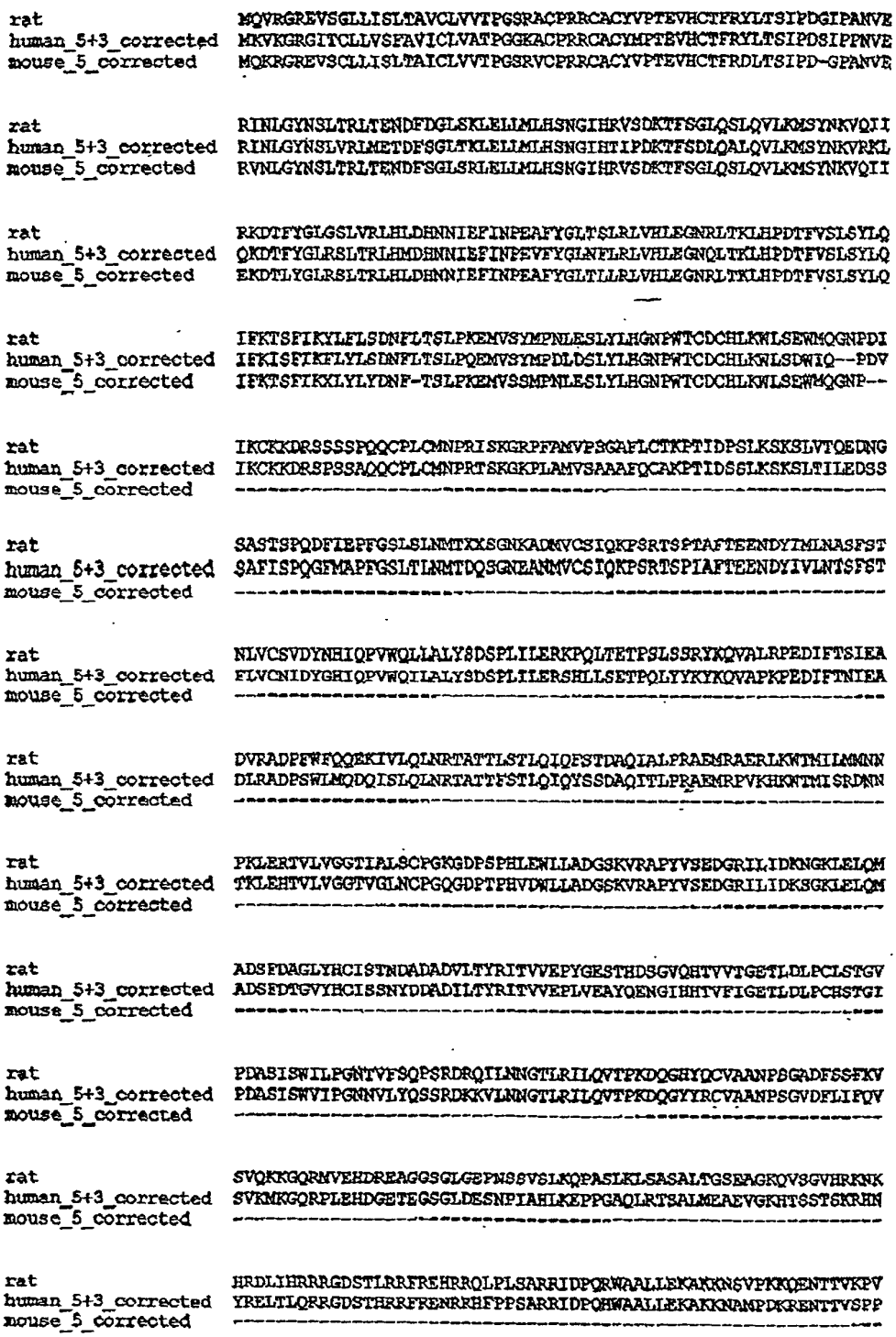

rat_cDNA
human_5+3_corrected
mus_cDNA_5
AAATTGAACAGAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTG
AAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAG
ACTGCAAGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAG
ATTGCAAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCA
TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTTT
TGATCAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTT
TCCACAATGGAACCTTGTATTTCAACAAGGTTGGGATGGCAGAGGAAGGAGATTATATCT
TCAACAATGGAACTTTATACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTT
GCTCTGCCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGTTCTAACAG
GCTATGCCCAGAACACCCTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAG
CCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGGCTGGAGAAACAG
CTGCTCCCCGGATAAGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAG
CTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTATTTTGGTTGCTGCCTT
CTGTCCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTT
CCAACAATGTCATTTCATTCTCCAATGACAGGTTCACATTTCATGCCAATAGAACTTTGT
CCAATGACATGATTTCCTTCTCCATTGATAGGTACACATTTCATGCCAATGGGTCTTTGA
CCATCCATAAAGTGAAACCACTTGACTCTGGGGACTATGTGTGCGTAGCTCAGAATCCTA
CCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCA
GTGGGGATGACACTAAGACATACAAACTGGACATTGTCTCTAAACCTCCATTAATCAATG
GTGGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATG
GCCTGTATGCAAACAAGACTGTTATTAAAGCCACAGCCATTCGGCACTCCAAAAAATACT
GTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACT
TTGACTGCAGAGCAGATGGGATCCCATCTTCCCAGGTCACGTGGATTATGCCAGGCAATA
TTGACTGCAGAGCTGAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATA
TTTTCCTCCCAGCTCCATACTTTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGG
TTTTCCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAATGGAACCTTGG
AGATGAGGAACATCCGGCTTTCTGACTCTGCGGACTTCACCTGTGTGGTTCGGAGCGAGG
AAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTGTGGCCCGAAATGAAG


rat_cDNA
human_5+3_corrected
mus_cDNA_5
GAGGAGAGAGTGTGTTGGTAGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACAT
GTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT
rat_cDNA
human_5+3_corrected
mus_cDNA_5
TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACT
TTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGAAAGTCCACAGCATTGAATT
rat_cDNA
human_5+3_corrected
mus_cDNA_5
GCTCTGTGGATGGGAACCCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGT
GCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGAT
rat_cDNA
human_5+3_corrected
mus_cDNA_5
TTGCTAACAGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTT
TTTCCAATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTT
rat_cDNA
human_5+3_corrected
mus_cDNA_5
ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT
CTAAAACAACTCGGGAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCT
rat_cDNA
human_5+3_corrected
mus_cDNA_5
ACATCGAGAAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGAAC
ATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCAC
rat_cDNA
human_5+3_corrected
mus_cDNA_5
CAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTGATGGGA
CAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTGTCTGATGGAA
rat_cDNA
human_5+3_corrected
mus_cDNA_5
TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCCTCAAG
TCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCCTCAAA
rat_cDNA
human_5+3_corrected
mus_cDNA_5
TGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAAAGCAACAACAGCTC
TTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCTT
rat_cDNA
human_5+3_corrected
mus_cDNA_5
ACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTGGCCAGGCAGTTATTAGCG
ATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCATACACTGATTACTG
rat_cDNA
human_5+3_corrected
mus_cDNA_5
TGTCAGTGATGGTTGTGGCCTACCCTCCCCGAATCATAAACTACCTACCCAGGAACATGC
TTCCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTG
rat_cDNA
human_5+3_corrected
mus_cDNA_5
TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCCTTGGGAATCCCCAAGCCAA
TCACCAGGACAGGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAG
rat_cDNA
human_5+3_corrected
mus_cDNA_5
AAGTCACCTGGGAGACGCCAAGACACTCCCTGCTCTCAAAAGCAACAGCAAGAAAACCCC
AAATCACATGGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACAC
rat_cDNA
human_5+3_corrected
mus_cDNA_5
ATAGAAGTGAGATGCTTCACCCACAAGGTACGCTGGTCATTCAGAATCTCCAAACCTCGG
ATGGAAGTGAGCAGCTTCACTTACAAGGTACCCTAGTCATTCAGAATCCCCAAACCTCCG



(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

Figure 15



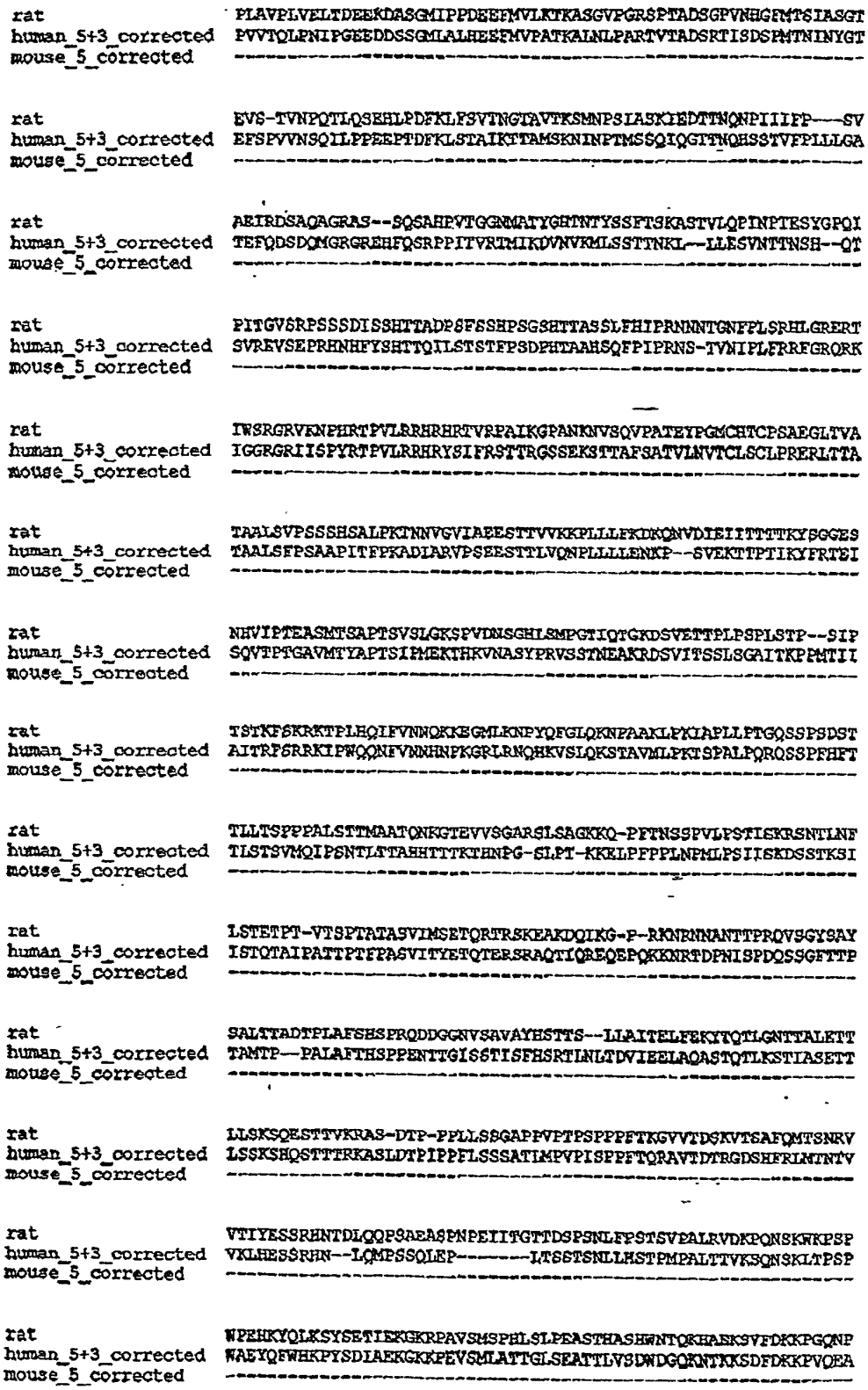
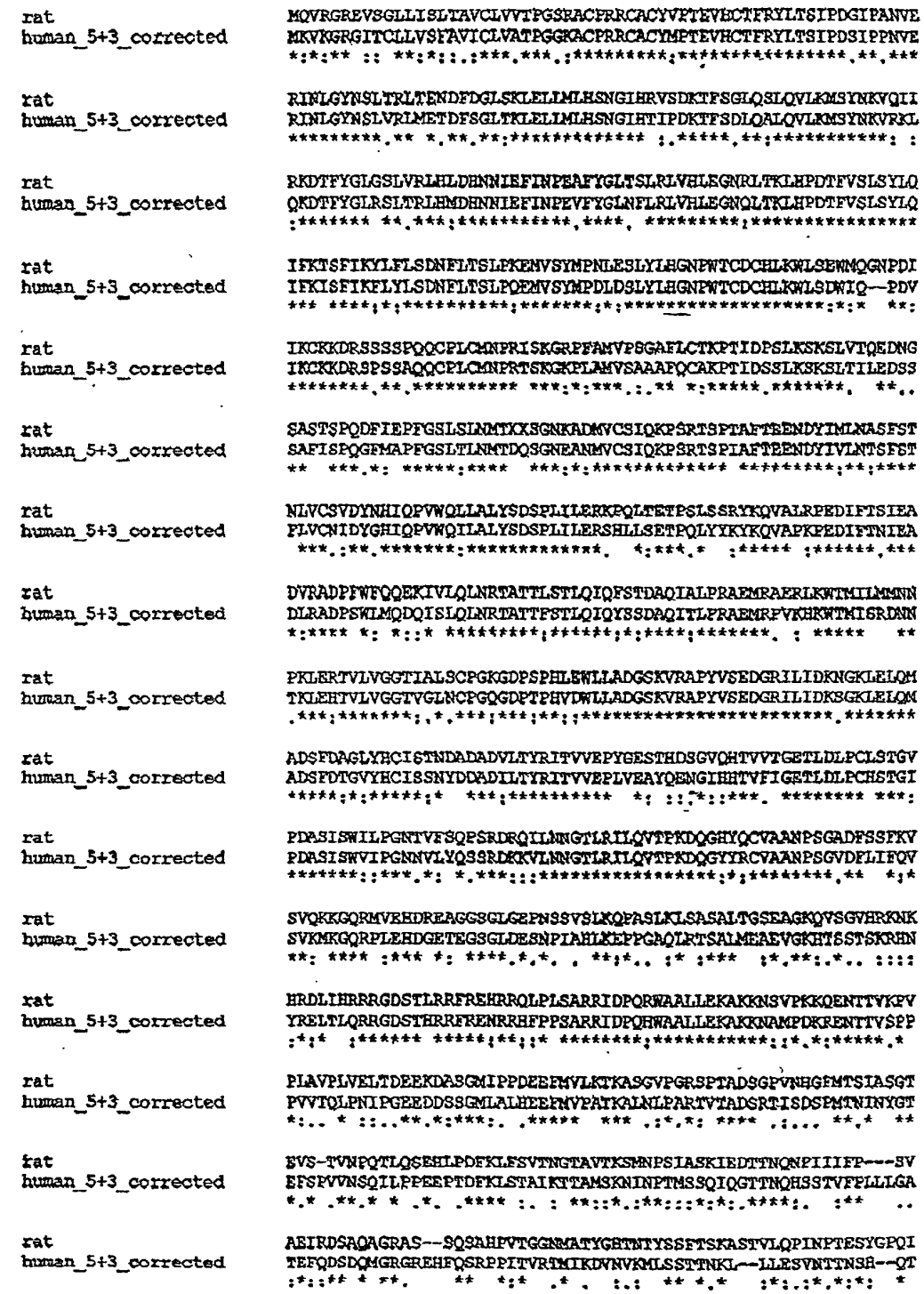
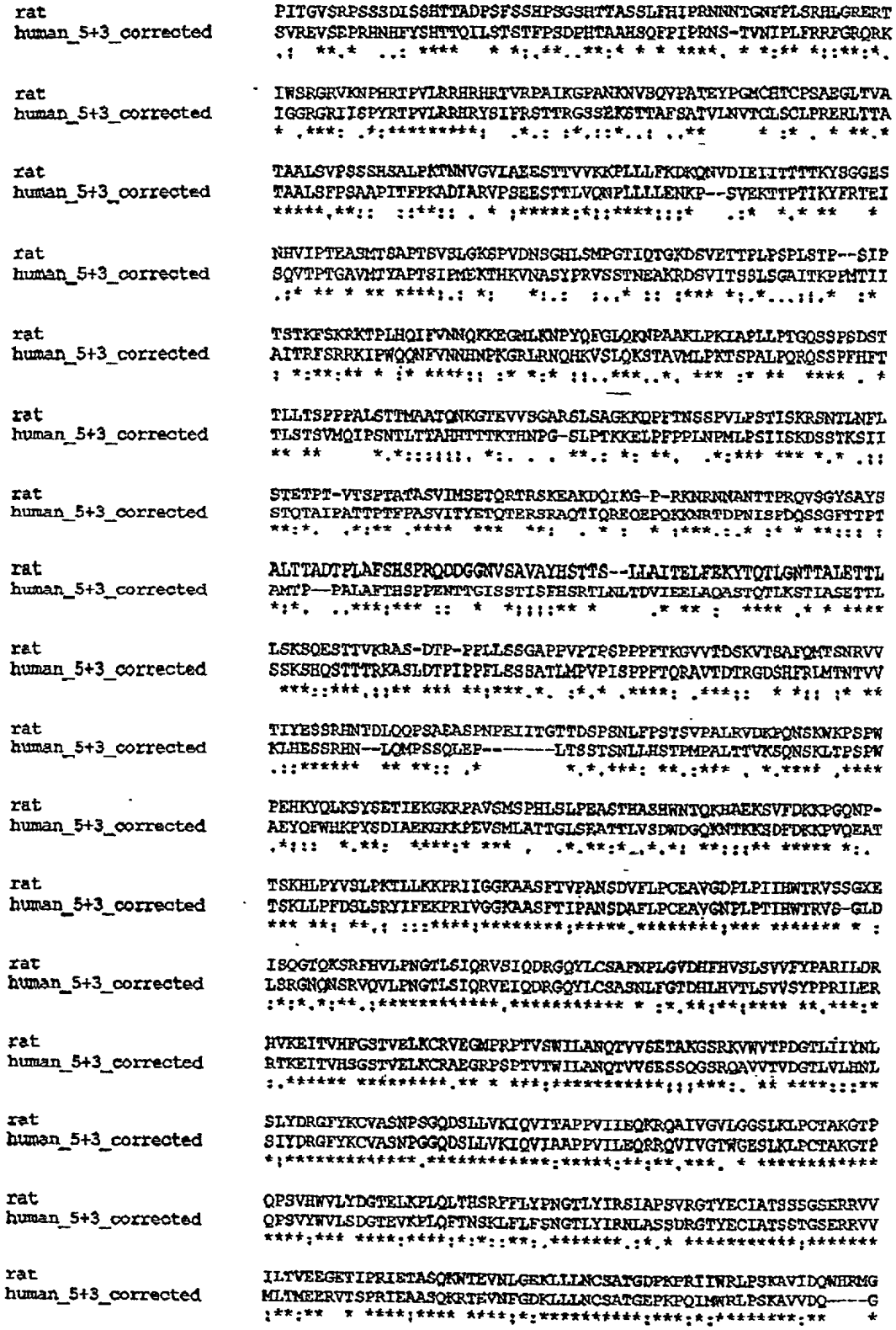
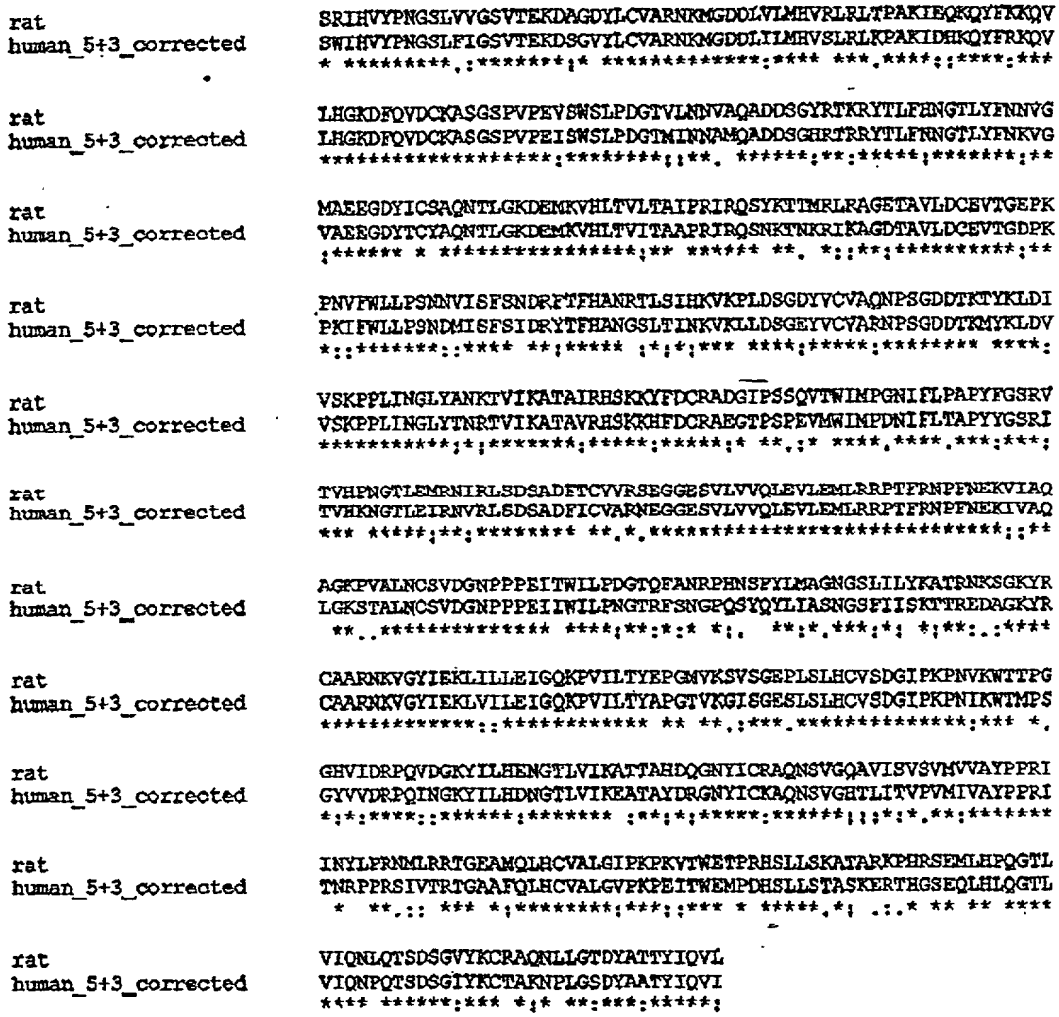

rat
human_5+3_corrected
mouse_5_corrected

Figure 16


rat
human_5+3_corrected
MQVRGREVSGLLISLTAVCLVVTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLTSIPDSIPPNVE
rat
human_5+3_corrected
RINLGYNSLTRLTENDFDGLSKLELIMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RINLGYNSLVRLMETDFSGLTKLELIMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRKL
rat
human_5+3_corrected
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
QKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYLQ
rat
human_5+3_corrected
IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSEWMQGNPDI
IFKISFIKFLYLSDNFLTSLPQEMVSYMPDLDSLYLHGNPWTCDCHLKWLSDWIQ—PDV
rat
human_5+3_corrected
IKCKKDRSSSSPQQCPLCMNPRISKGRPFANVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
IKCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAPQCAKPTIDSSLKSKSLTILEDSS
rat
human_5+3_corrected
SASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
SAFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFST
rat
human_5+3_corrected
NLVCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
PLVCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEA
rat
human_5+3_corrected
DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMIIMMNN
DLRADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEMRFVKHKWTMISRDNN
rat
human_5+3_corrected
PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQM
TKLEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
rat
human_5+3_corrected
ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGI
rat
human_5+3_corrected
PDASISWILPGNTVFSQPSRDRQILMNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
PDASISWVIPGNNVLVQSSRDKKVLNNGTLRILQVTPKDQGYYRCVAANPSGVDFLIFQV
rat
human_5+3_corrected
SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSASALTGSEAGKQVSGVHRKNK
SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRHN
rat
human_5+3_corrected
HRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENITVKPV
YRELTLQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALLEKAKKNAMPDKRENTTVSPP
rat
human_5+3_corrected
PLAVPLVELTDEEKDASGMIPPDEEPMVLKTKASGVPGRSPTADSGPVNHGPMTSIASGT
PVVTQLPNIPGEEDDSSGMLALHEEPMVPATKALNLPARTVTADSRTISDSPMTNINYGT
rat
human_5+3_corrected
EVS-TVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIIFP---SV
EFSPVVNSQILPPEEPTDFKLSTAIKTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLGA
rat
human_5+3_corrected
AEIRDSAQAGRAS--SQSAHPVTGGNMATYGHTNTYSSFTSKASTVLQPINPTESYGPQI
TEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKL--LLESVNTTNSH--QT


rat
human_5+3_corrected
PITGVSRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERT
SVREVSEPRHNHFYSHTTQILSTSTFPSDPHTAAHSQFPIPRNS-TVNIPLFRRFGRQRK
rat
human_5+3_corrected
IWSRGRVKNPHRTPVLRRHRHRTVRPAIKGPANKNVSQVPATEYPGMCHTCPSAEGLTVA
IGGRGRIISPYRTPVLRRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRERLTTA
rat
human_5+3_corrected
TAALSVPSSSHSALPKTNNVGVIAEESTTVVKKPLLLFKDKQNVDIEIITTTTKYSGGES
TAALSFPSAAPITFPKADIARVPSEESTTLVQNPLLLLENKP--SVEKTTPTIKYFRTEI
rat
human_5+3_corrected
NHVIPTEASMTSAPTSVSLGKSPVDNSGHLSMPGTIQTGKDSVETTPLPSPLSTP--SIP
SQVTPTGAVMTYAPTSIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLSGAITKPPMTII
rat
human_5+3_corrected
TSTKFSKRKTPLHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDST
AITRFSRRKIPWQQNFVNNHNPKGRLRNQHKVSLQKSTAVMLPATSPALPQRQSSPFHFT
rat
human_5+3_corrected
TLLTSPPPALSTPMAATQNKGTEVVSGARSLSAGKKQPFTNSSPVLPSTISKRSNTLNFL
TLSTSVMQIPSNTLTTAHTTTKTHNPG-SLPTKKELPFPPLNPMLPSIISKDSSTKSII
rat
human_5+3_corrected
STETPT-VTSPTATASVIMSETQRTRSKEAKDQIKG-P-RKNRNNANTTPRQVSGYSAYS
STQTAIPATTPTFPASVITYETQTERSRAQTIQREQEPQKKNRTDPNISPDQSSGFTTPT
rat
human_5+3_corrected
ALTTADTPLAFSHSPRQDDGGNVSAVAYHSTTS--LLAITELFEKYTQTLGNTTALETTL
AMTP---PALAFTHSPPENTTGISSTISFHSRTLNLTDVIEELAQASTQTLKSTIASETTL
rat
human_5+3_corrected
LSKSQESTTVKRAS-DTP-PPLLSSGAPPVPTRSPPPFTKGVVTDSKVTSAFQMTSNRVV
SSKSHQSTTTRKASLDTPIPPFLSSSATLMPVPISPPFTQRAVTDTRGDSHFRLMTNTVV
rat
human_5+3_corrected
TIYESSRHNTDLQQPSAEASPNPEIITGTTDSPSNLFPSTSVPALKVDKPQNSKWKPSPW
KLHESSRHN--LQMPSSQLEP--------LTSSTSNLLHSTPMPALTTVKSQNSKLTPSPW
rat
human_5+3_corrected
PEHKYQLKSYSETIEKGKRPAVSMSPHLSLPEASTHASHWNTQKHAEKSVFDKKPGQNP-
AEYQFWHKPYSDIAEKGKKPEVSMLATTGLSEATTLVSDWDGQKNTKKSDFDKKPVQEAT
rat
human_5+3_corrected
TSKHLPYVSLPKTLLKKPRIIGGKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGKE
TSKLLPFDSLSRYIFEKPRIVGGKAASFTIPANSDAFLPCEAVGNPLPTIHWTRVS-GLD
rat
human_5+3_corrected
ISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDR
LSRGNQNSRVQVLPNGTLSIQRVEIQDRGQYLCSASNLFGTDHLHVTLSVVSYPPRILER
rat
human_5+3_corrected
HVKEITVHPGSTVELKCRVEGMPRPTVSWILANQTVVSETAKGSRKVWVTPDGTLIIYNL
RTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSESSQGSRQAVVTVDGTLVLHNL
rat
human_5+3_corrected
SLYDRGFYKCVASNPSGQDSLLVKIQVITAPPVIIEQKRQAIVGVLGGSLKLPCTAKGTP
SIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRRQVIVGTWGESLKLPCTAKGTP
rat
human_5+3_corrected
QPSVHWVLYDGTELKPLQLTHSRFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVV
QPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLASSDRGTYECIATSSTGSERRVV
rat
human_5+3_corrected
ILTVEEGETIPRIETASQKNTEVNLGEKLLLNCSATGDPKPRIIWRLPSKAVIDQNHRMG
MLTMEERVTSPRIEAASQKRTEVNFGDKLLLNCSATGEPKPQIMWRLPSKAVVDQ----G



(rat: SEQ ID NO:13)
(human_5+3_corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYVPTEVHCTFRDLTSIPDGPANVER
VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
KTSFIKXLYLYDNFTSLPKEMVSSMPNLESLYLHGNPWTCDCHLKWLSEWMQGNP
(SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYMPTEVHCTFRYLTSIPDSIPPNVE
RINLGYNSLVRLMETDFSGLTKLELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
QIFKISFIKFLYLSDNFLTSLPQEMSYMPDLDSLLYLHGNPWTCDCHLRWLSDWIQPDVI
KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
VCNIDYGHIQPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
ADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEMRPVKHKWTMISRDNNTK
LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFIGETLDLPCHSTGIP
DASISWVIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDQGYYRCVAANPSGVDFLIFQV
SVKMKGQRPLEHDGETEGSGLDESNPIAHLKEPPGAQLRTSALMEAEVGKHTSSTSKRH
NYRELTLQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALLEKAKKNAMPDKRENTTV
SPPPVVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
YGTEFSPVVNSQILPPEEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG
ATEFQDSDQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
SEPRHNHFYSHTTQILSTSTFPSDPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS
PYRTPVLRRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
SIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLSGAITKPPMTIIAITRFSRRKIPWQQNFV
NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNTLTT
AHHTTTKTHNPGSLPTKKELPFPPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
TQTERSRAQTIQREQEPQKKNRTDPNISPDQSSGFTTPTAMTPPALAFTHSPPENTTGISST
ISFHSRTLNLTDVIEELAQASTQTLKSTIASETTLSSKSHQSTTTRKASLDTPIPPFLSSSAT
LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLQMPSSQLEPLTSSTSNLL
HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGKKPEVSMLATTGLSEATTLV
SDWDGQKNTKKSDFDKKPVQEATTSKLLPFDSLSRYIFEKPRIVGGKAASFTIPANSDAF
LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCSASNL
FGTDHLHVTLSVVSYPPRILERRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSE
SSQGSRQAVVTVDGTLVLHNLSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRR
QVIVGTWGESLKLPCTAKGTPQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPRIEAASQKRTEVNFGDKLLLNCSATGEP
KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
LKPAKIDHKQYFRKQVLHGKDFQVDCKASGSPVPEISWSLPDGTMINNAMQADDSGHR
TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTN
KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
EYVCVARNPSGDDTKMYKLDVVSKPPLINGLYTNRTVIKATAVRHSKKHFDCRAEGTP
SPEVMWIMPDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGESVLVVQLE
VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEIIWILPNGTRFSNGPQSY
QYLIASNGSFIISKTTREDAGKYRCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGE
SLSLHCVSDGIPKPNIKWTMPSGYVVDRPQINGKYILHDNGTLVIKEATAYDRGNYICKA
QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGAAFQLHCVALGVPKPEITWEMPDHS
LLSTASKERTHGSEQLHLQGTLVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
NO: 16)

Figure 19

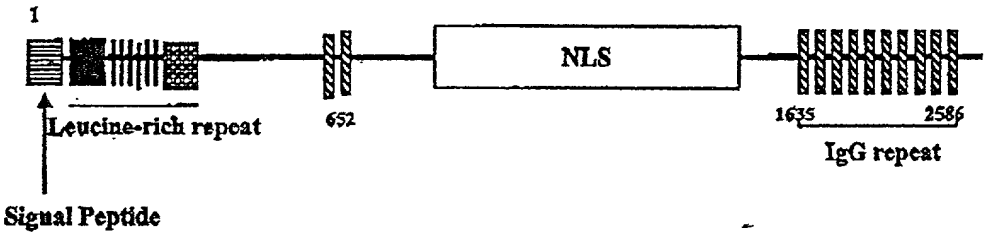
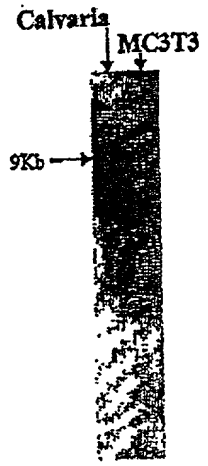


Figure 20

A



B

Cell lines & primary cells		
C6	Rat glioma	-
C2C12	Myoblasts	-
STO	Fibroblasts	-
Bone marrow	Stromal cells	- (+)
PC-3	prostatic carcinoma	-
SAOS-2	osteosarcoma	-
U2OS	osteosarcoma	+
Articular cartilage	chondrocytes	-
ROS	osteosarcoma	-
Calvaria	Osteoblasts	+
C3H10T1/2	mesenckyme	- (+)
MC3T3-E1	preosteoblastic	- (+)

Figure 21

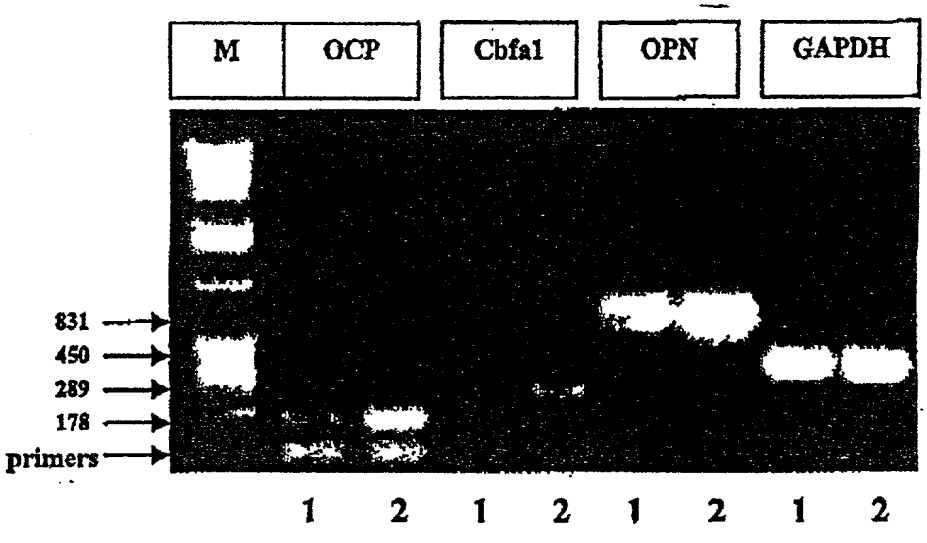


Figure 22

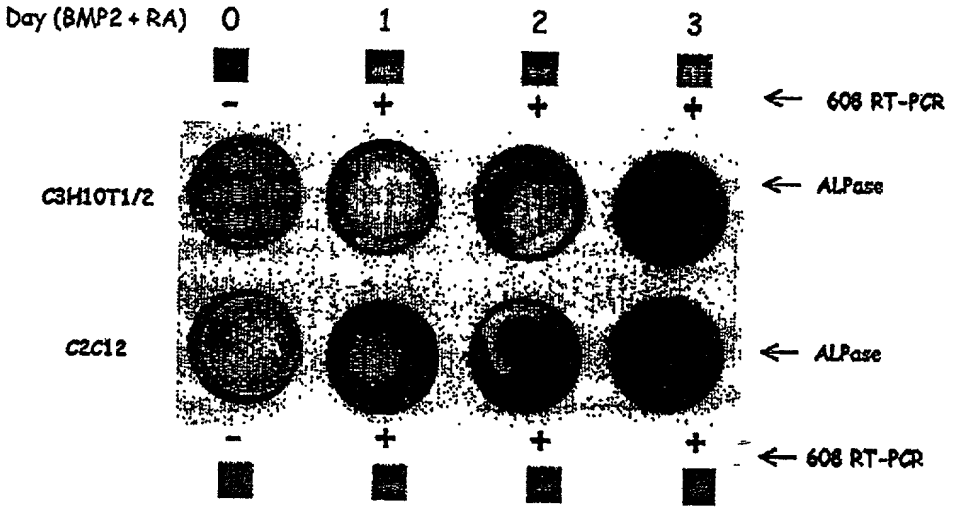


Figure 23

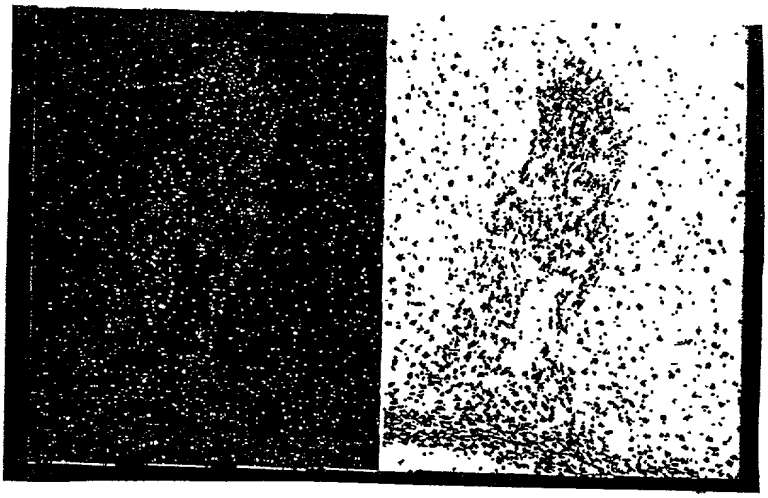


Figure 24

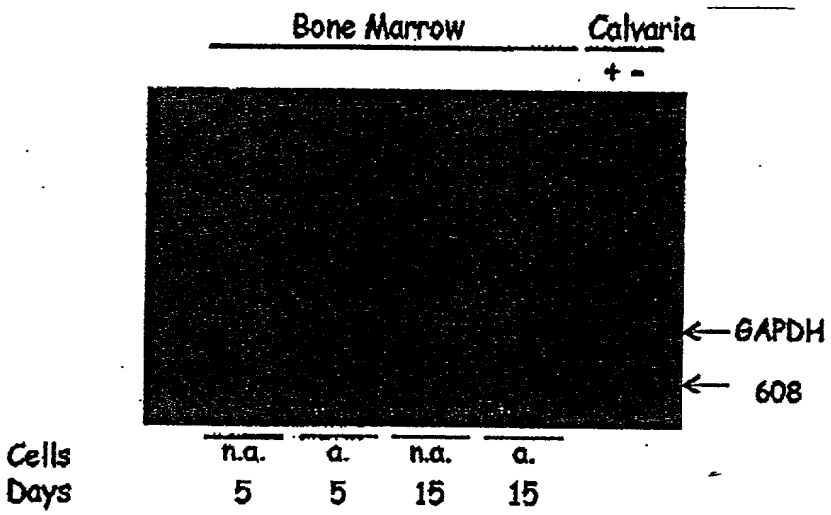


Figure 25

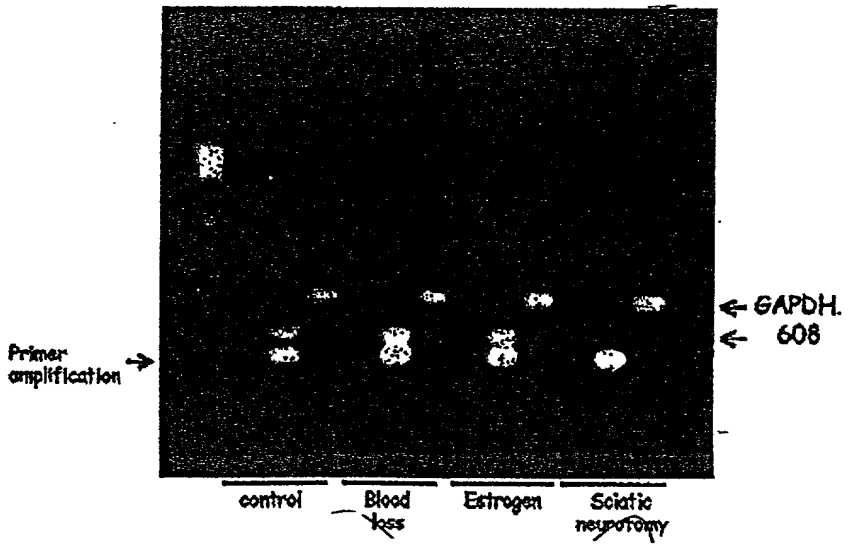


Figure 26

Figure 27

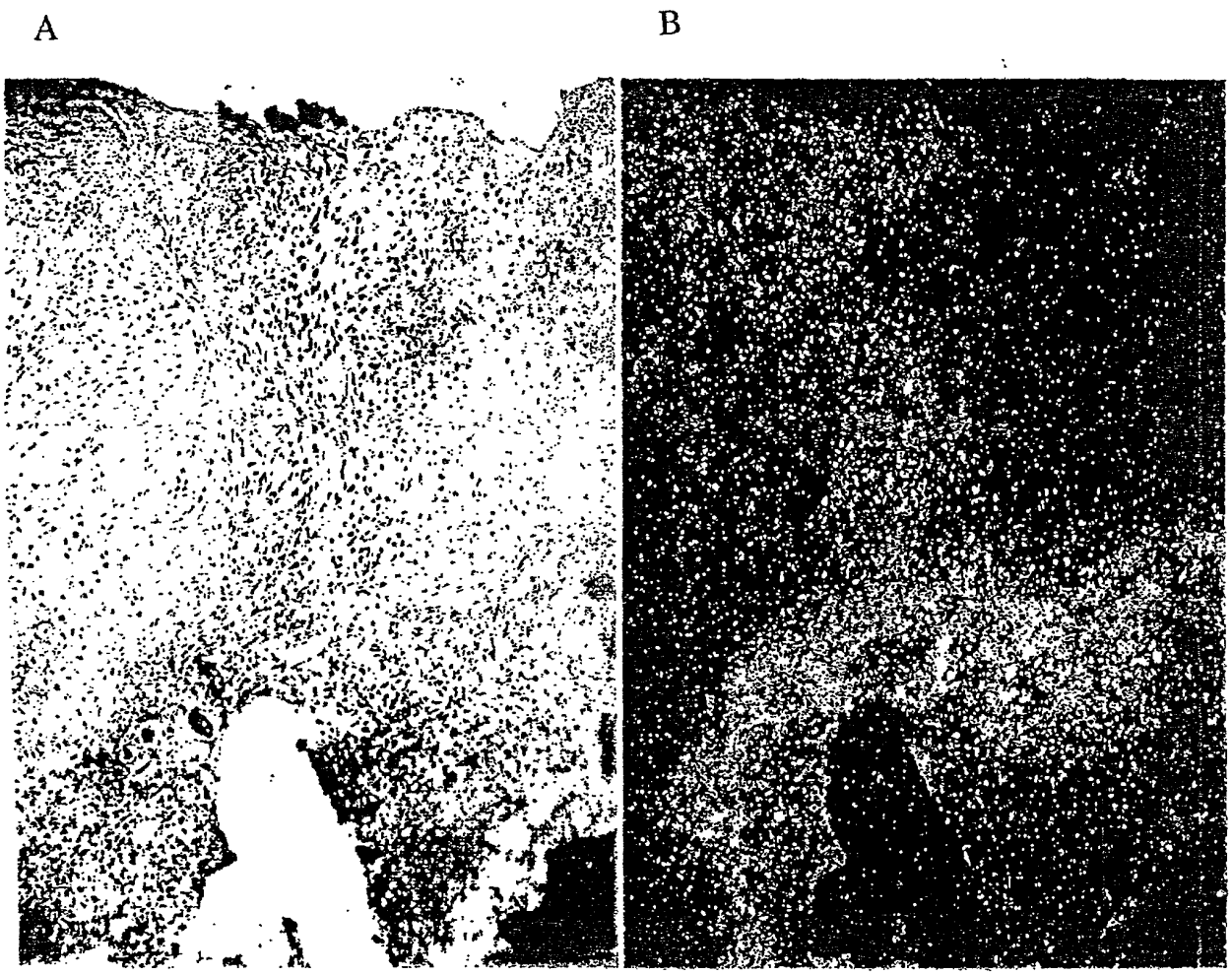


Figure 28

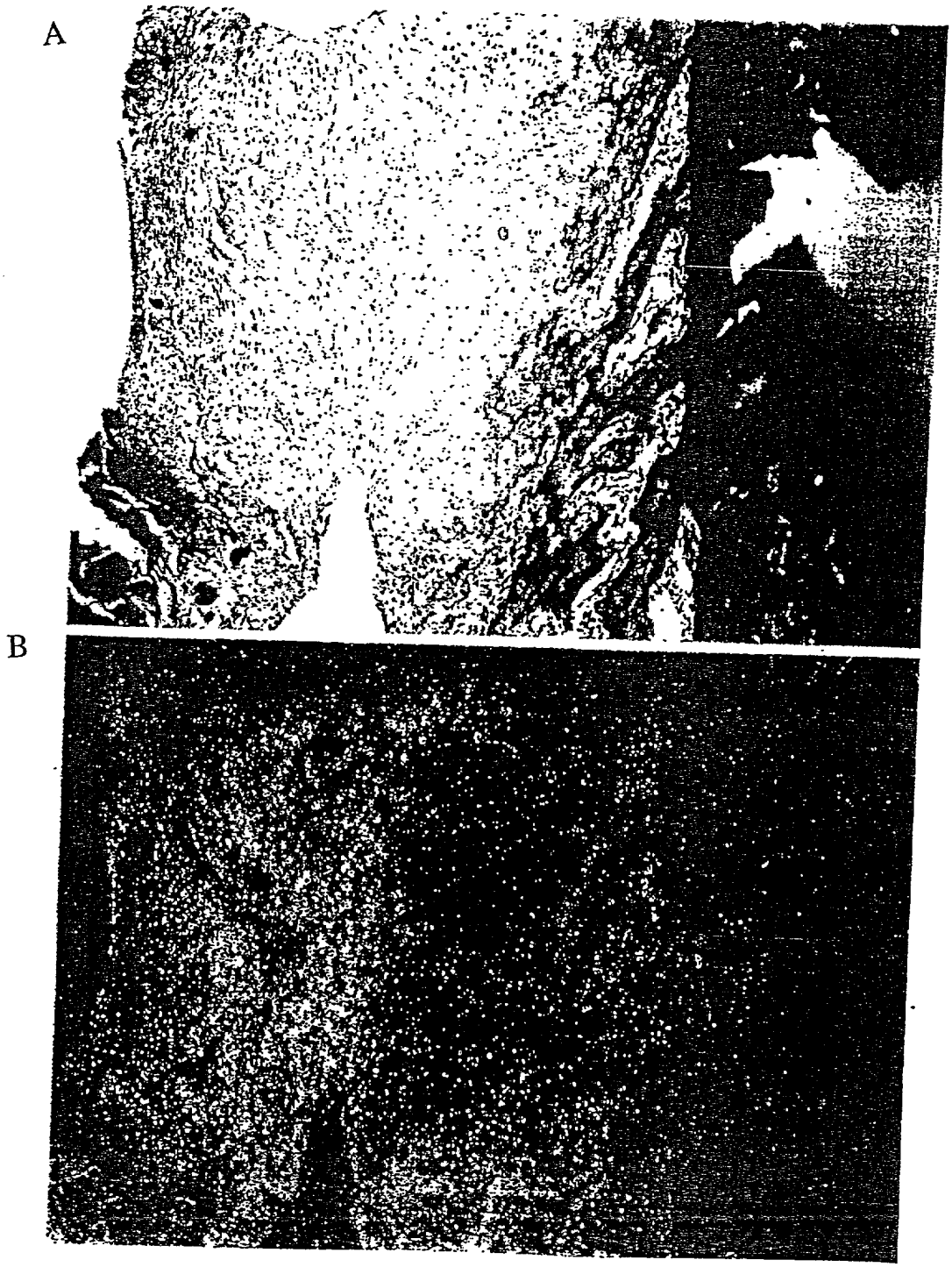


Figure 29

A

B

Figure 30

Figure 31

Figure 32

A

B

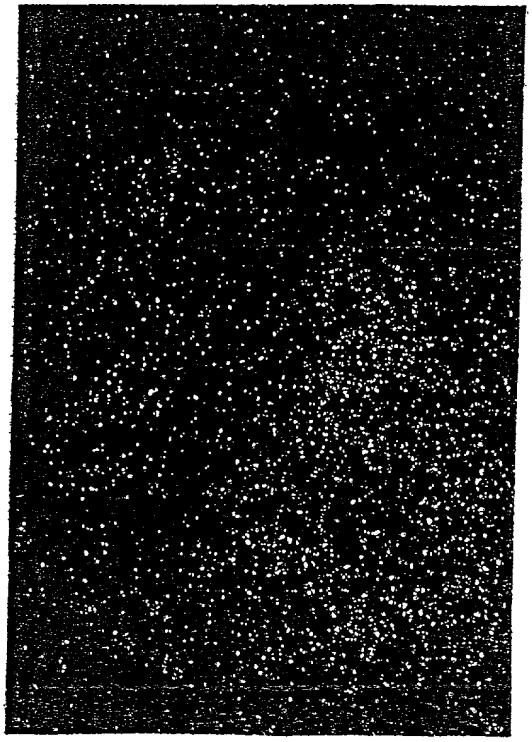


Figure 33

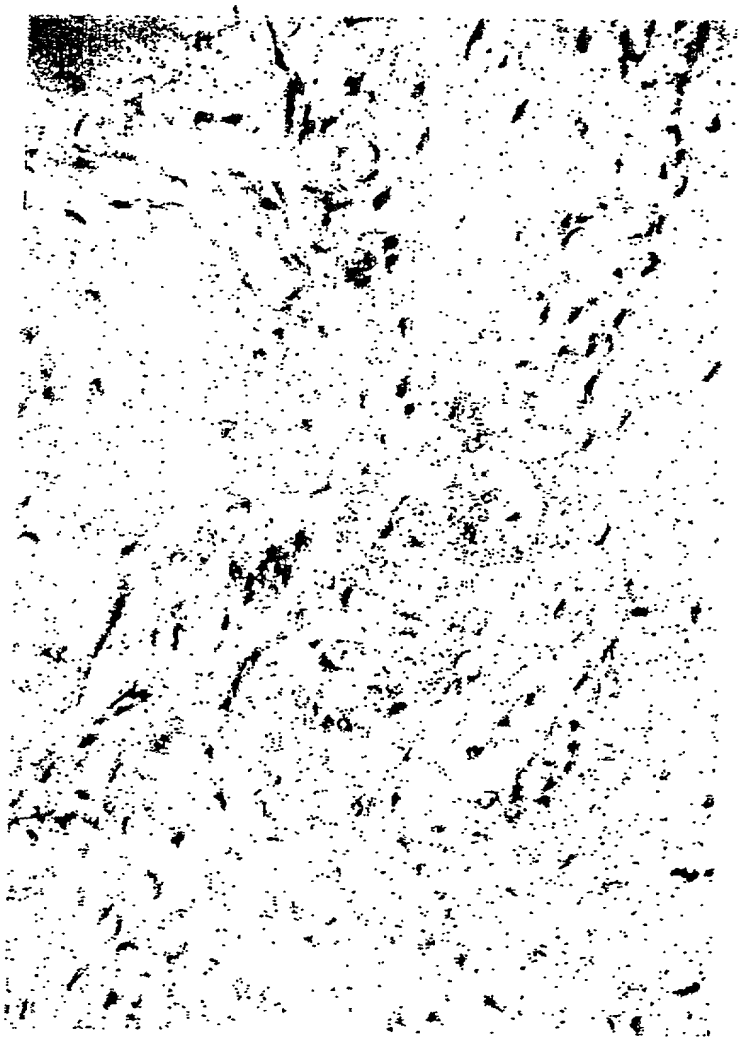


Figure 34

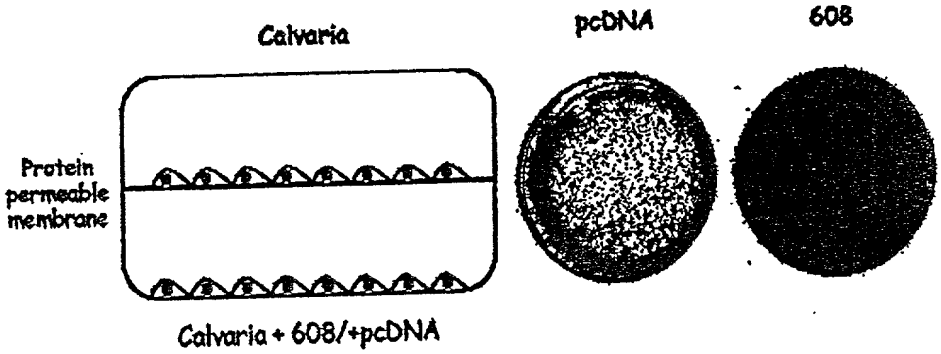


Figure 35

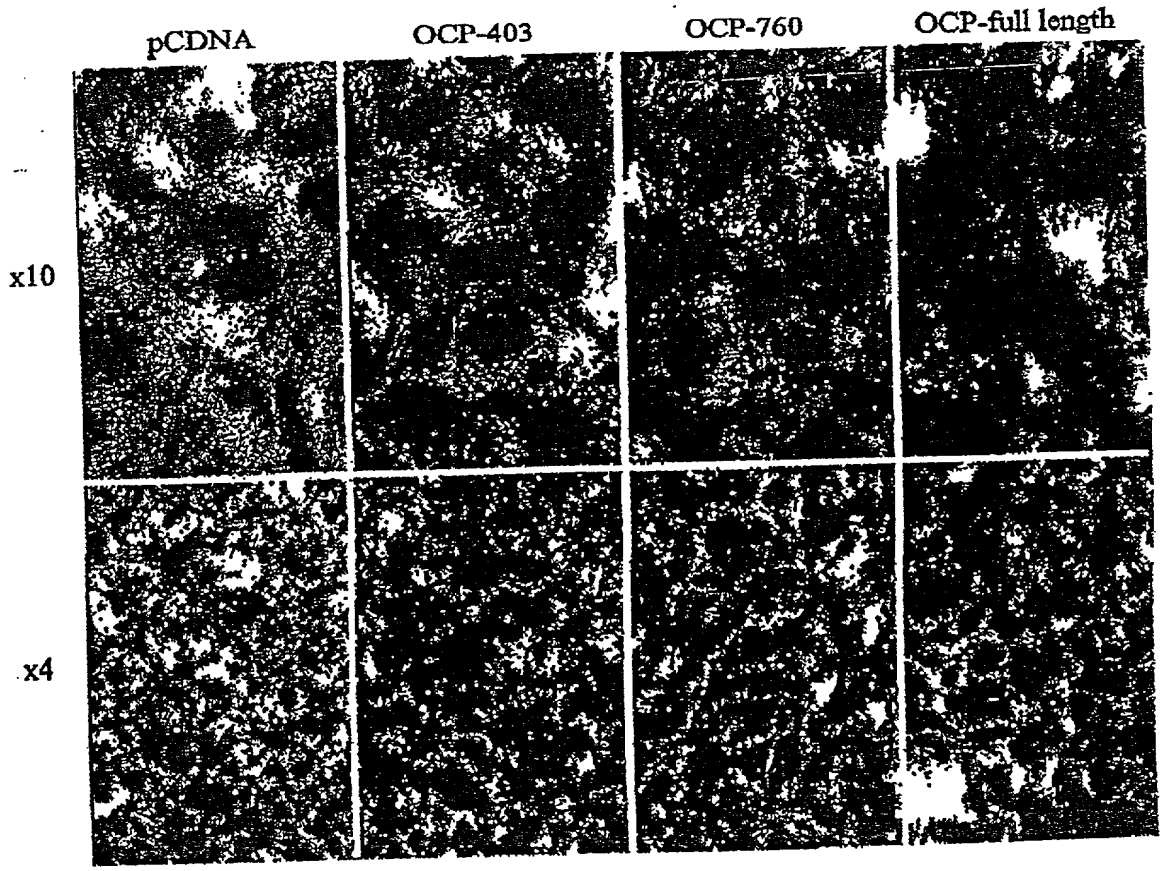


Figure 36

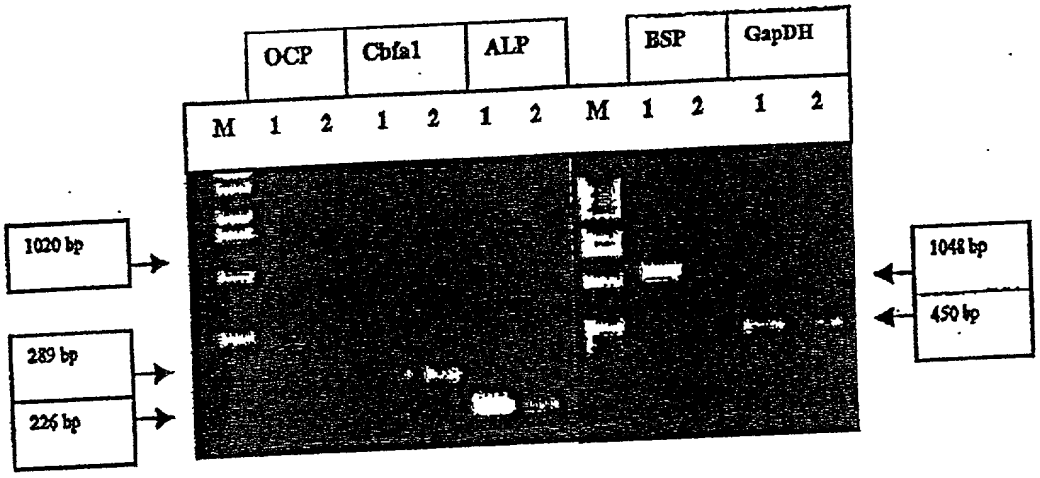


Figure 37

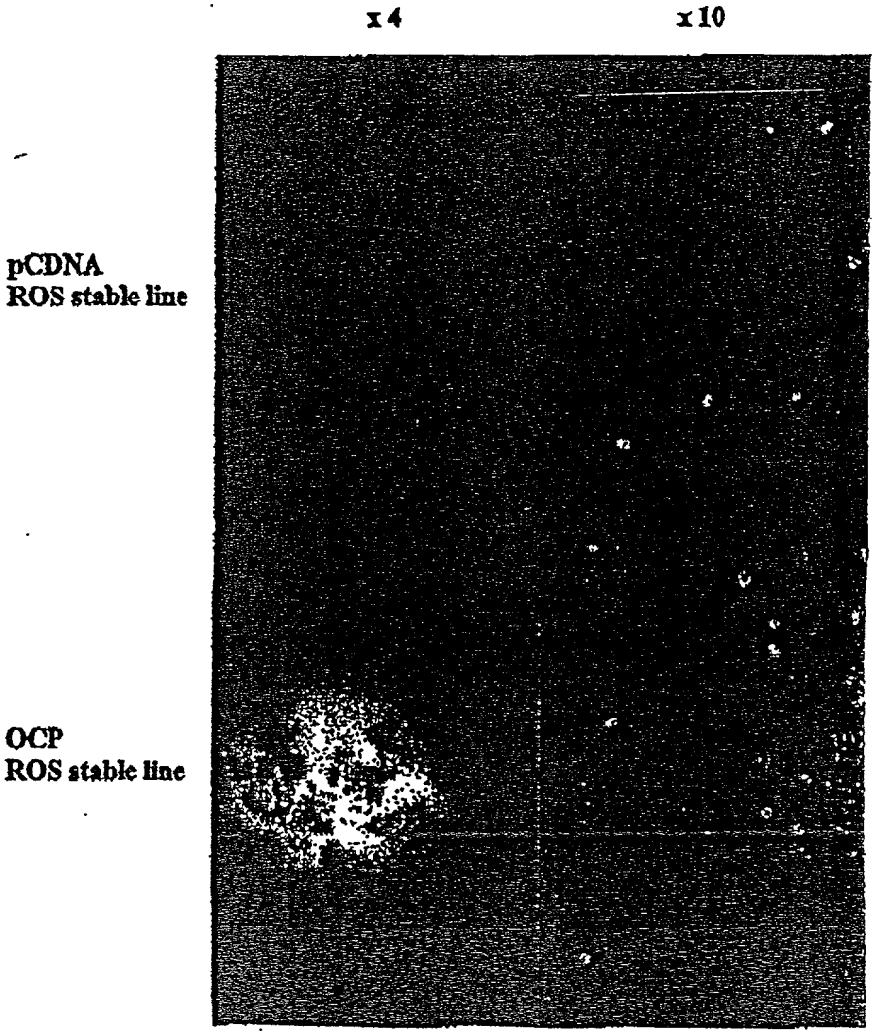


Figure 38

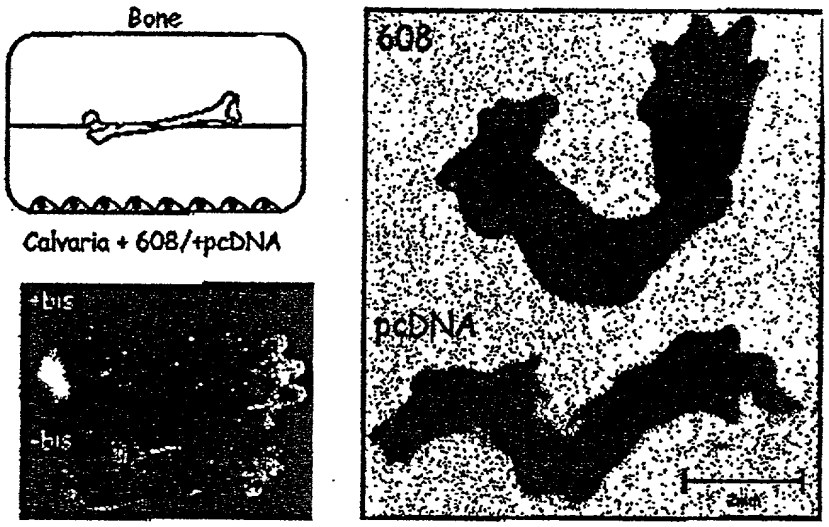


Figure 39

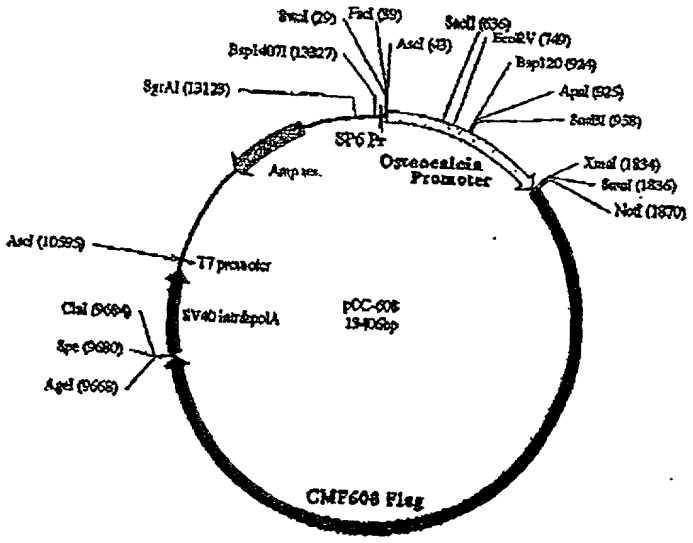


Figure 40

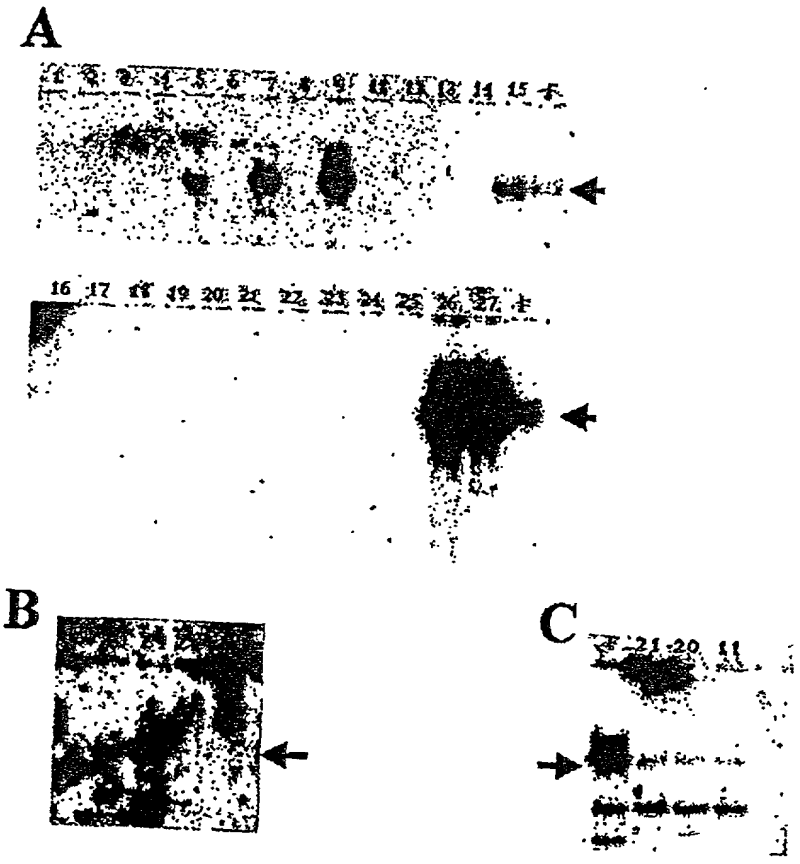


Figure 41

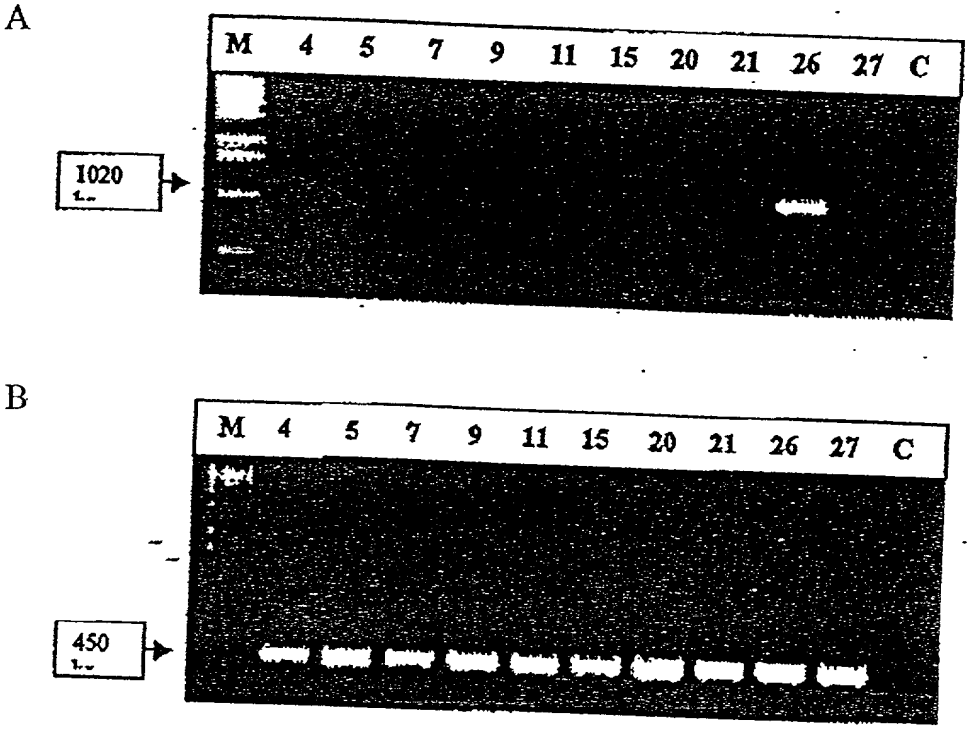
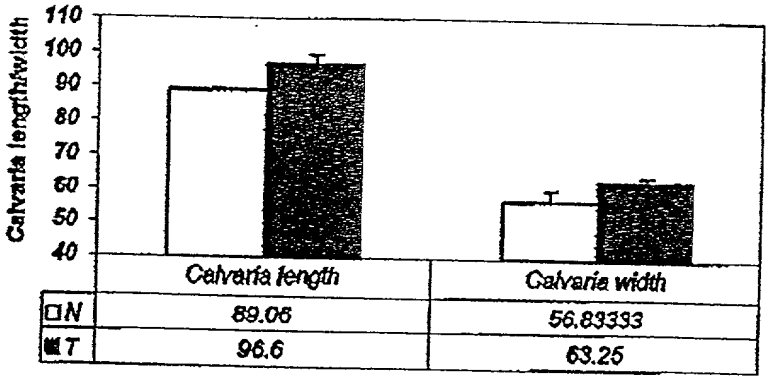


Figure 42



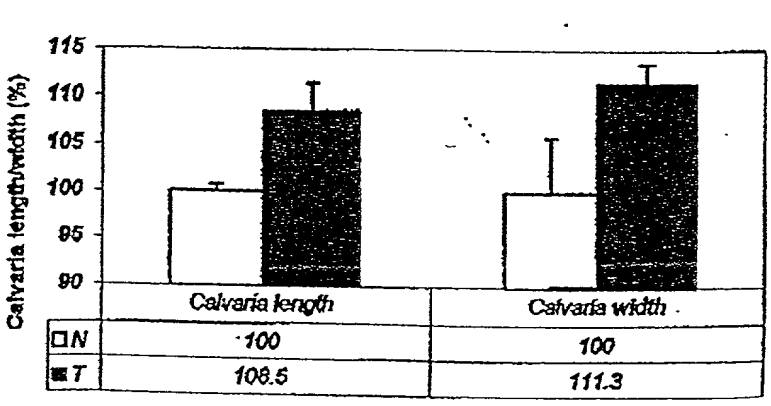


Figure 43

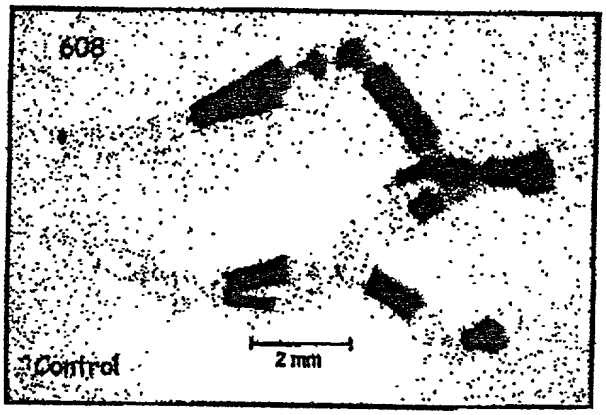


Figure 44

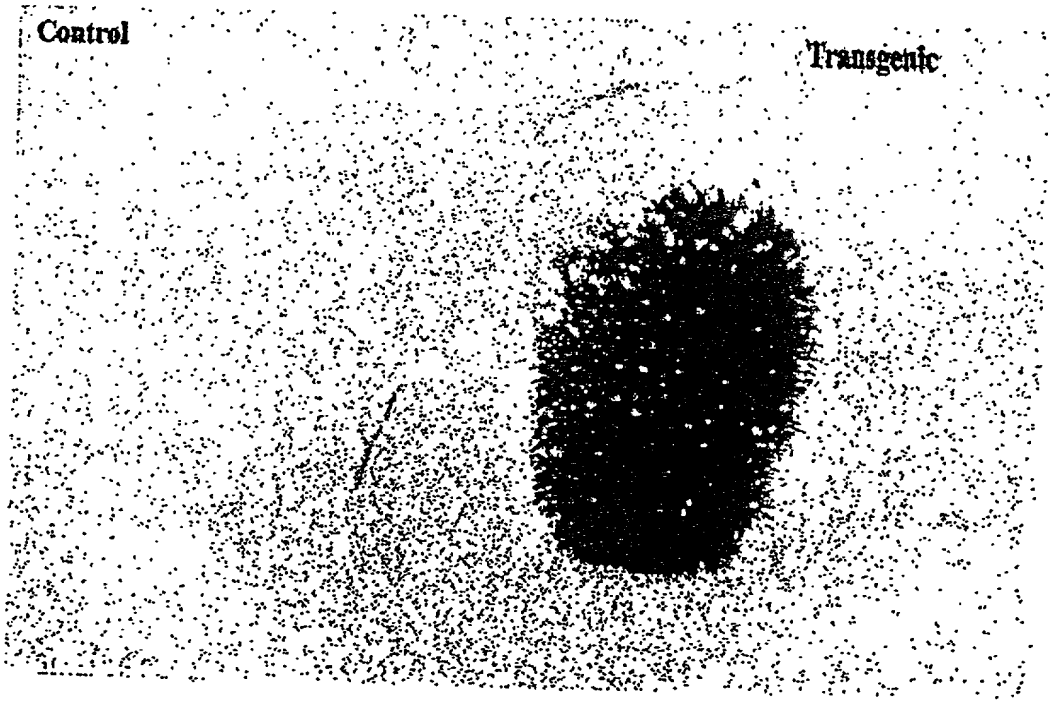


Figure 45

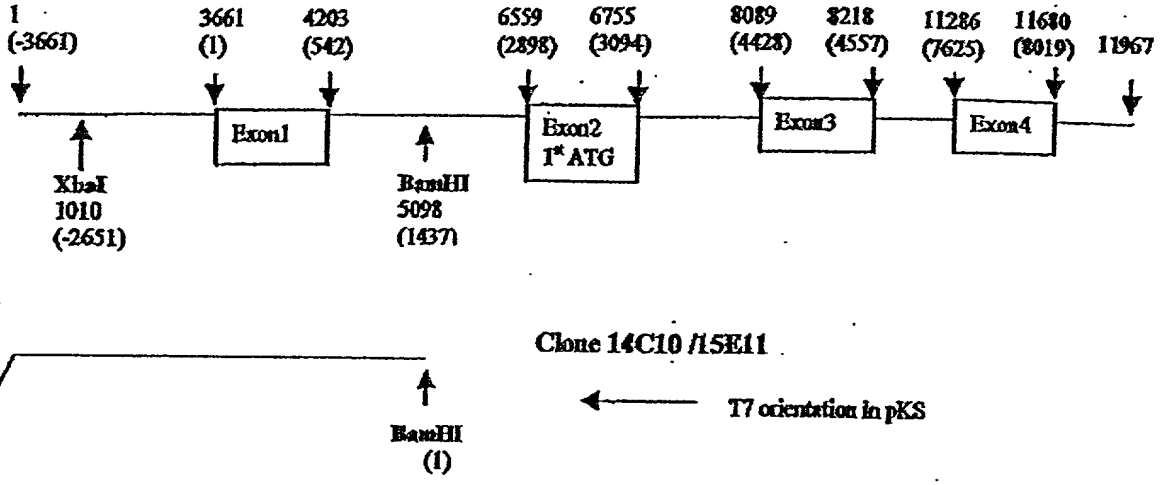


Figure 46

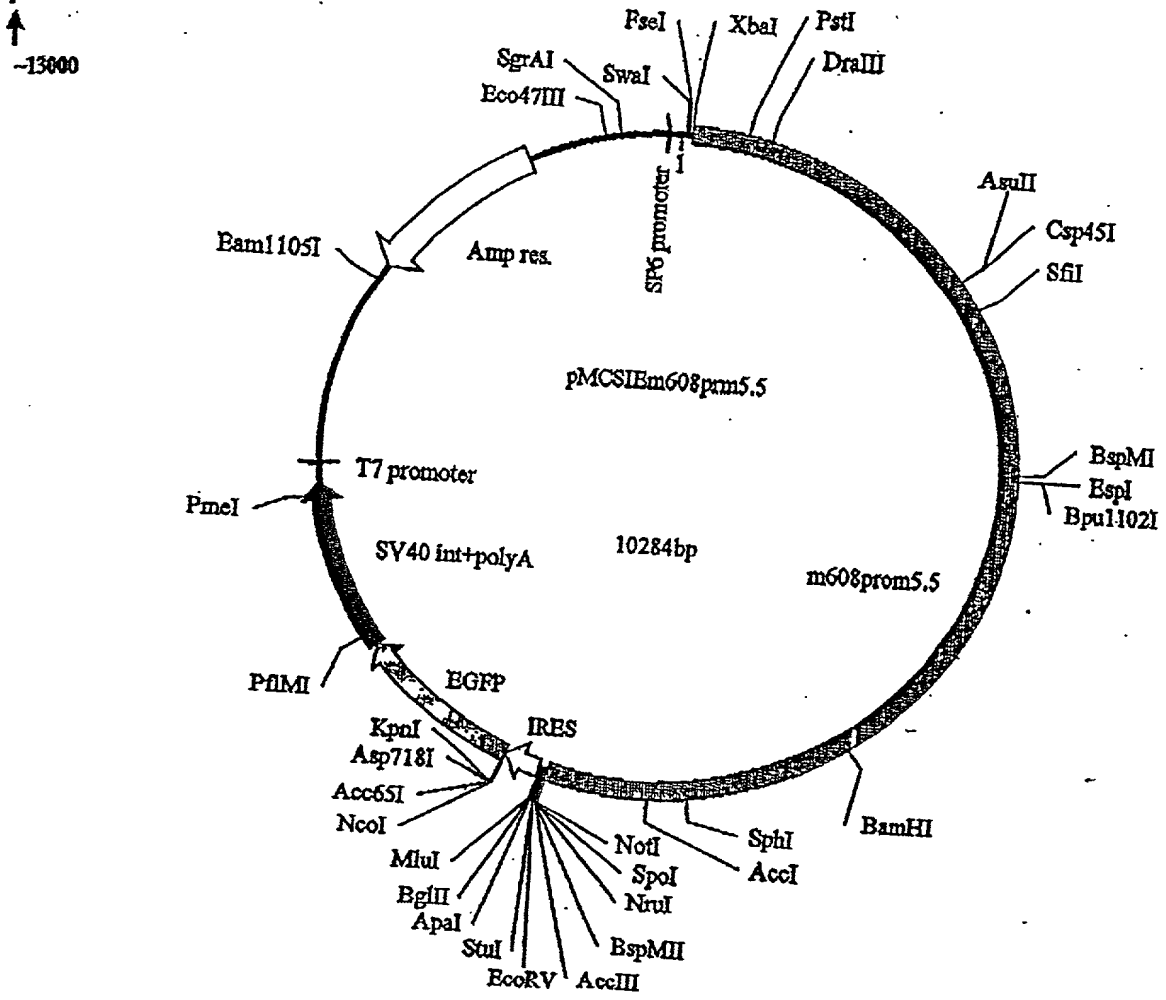


Figure 47

AGATTTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAGG
TATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATGTGGTA
TTCATTTACGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTTAATCATC
ACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACAGTGTAGGCATAGAAGTA
ATCATTAGTAAATCTTAATTTGGGTTAAACTCATTCATAACAGCTCCAGGTTGGGAGGGAT
CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCTTC
CTTGGCCATGCTCCCCATCACTGTGTCAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAG
AGCAGACAGTGTTATGCCTGCAAAGTCAGAGACTCAGCCCTCCCAGCTGGTCAGTTTACT
GTCCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGC
AGACATGCTAGAAAGAAATTTGACCTTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTAT
TTTGGAAGGTTAAGCCAAATAAATAAAACACATATTAAATAATACAATGTTACAAAAATT
GATCATATAAAGAAGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTAT
TTACTGGTGCAAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAATATTTTT
AAAATGTATATAAAAGTGTAAAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACT
(BamHI)
TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCAATG
GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAAGTTTACA
CTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTTGTTGTCATCATTTATC
TTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCATTTATTGGATAA
TTGGGGGATGGTGTTAATTTTTTTGCAGTTCTTATGGAATTCCAAAAAACAAAAAACAAAC
AAACAAACAAAAAACCTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
AGAAATCTGCACAGAATTTATTGCTACATTGTTCATTATTCACGACAGCCAAGAATGTGGA
ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAATGTGGAAATGTGTACAAC
AGAGTCCCATGTGGCCATAAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
GGAAATCAATTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATT
TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
AGCTAGCTGGGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
GGCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
TTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAGGGGCGTTCCAGTTTT
CTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGCATGATGGATGG
TTTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTGGTGGTTGTTGTGGTGGTG
GCTGTGGTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAACATGAAATTGGGTGGGTAGGA
AGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAGAAAAATTGTATGGAGCATATTTAA
ACAAACAAACAAACAAACAAAAGGTTCATTTTGCCACAAAAAGGTGTGAATTAAATTAAC
CAGTTACGACTCTTAAAGAAAATATTCCCAATTATTCCCAGAGTTGCTATGTATGCTGTGC
CTAGGACTTTGCTTGAACTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG – Artificial
NotI site

(SEQ ID NO:17)

Figure 48

CACAGACCTTCCTCTTCTAACCTCTCTCCCCCATCTTGTTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAACTGCTCCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTCTGACCACATCCCCAAGTTATCACAAGGAGTTT
TTCCTCAAACCTTTCCTCTCCAGCAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCCT
TCATTGGGAACCCTGTGCTCCATCCAATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT
CTTGGTCAGTGCAGTCTTTTGTATCCTCAAGAACACTGGGTCTGAAAATTTTAACCCAAAG
AACTGTTTTTTGTTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTTGTCTCCTCCTTTCAGTACCACCCTGCCTTTGCTGCCTTTCTCAAAGAATCAATAAA
ACCAAAGTGATATAGATTCATGGCATTCCTCTAACTGCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAGCCTGGAAGCAGTCACGAGGCAGTTTCACAGAAACTTAGCCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATGCCCCAGATTTGTCCCTGAGATATTGTGGTG
GGTCTTGCATGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAATCTTCATGACAGTCAGGTATTGCAAACTACATTGCATATTA
GAAGAAAGTTGGTAAATTCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAAGAAC
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAGGGGGGGGGGGTACT
GCTTTATTCATGATAAGGTCTGCTAGAGCAGAACCCCCTGGTGCTAGCTTTCACAAGGTTC
AAAGGTGTAGCATAAATTGTGACTAGAGTGTGAAATCTTTACCTGTCATTAGCTGACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCCTGTCAGTCCTTTGAAGGCAT
TCCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCCTGCTGTGACACCCTACTCCTTTGT
TTTCTGTATTATATAAGCCTGGTGTTCCCTTTGTGAAAAATTACATCCAGATACAGCACTC
CCTTGTGTCTGTGTCCTTTTGTCATTTCTGGCCAACTCCATGCCCACCTGCCAGAACCCCTA
GTCTTTTCCACAGATTGAGGGAGGCCGACTGAGCCTGGTCCATGGCATCTAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGCTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAACCCATCTCTTCCTCGTTTGCTAATATAGCAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTCCTGTATGTG
TGTCCATTGTTTCTTCAGTTCCTGTTGCCCCATTAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGCGTTTCAACGTTTGCCTG
TAAACAAGATATATGGTGCGCAGTTTATAATCATAATAAGCTTTGAAATAATATATAACTC
CATTCTCATTCTGCTTCCACGCTGAGCATCCTGTTTCCCCAGGGACCACAAGAGCATTTGA
AAAGTAGTGATTTATGACCTGCTTTGTTCTGTTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACTTGGGGTGACCTGTATCTGTGTTCCTGGGTCGTGCTCACTCATA
TTTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCACCATTCTACTTCCTGTAATCTGTATATGGTAGAAAAAGTTAATTT
ATGTGATTCTTCCAACTCCAAATATTTCAAATCTTTTAGCCCCTCAGCCTGGGATTTCTTTG
ACTAAGTCTATTGATTTGGAAGATCTCAGTGGTTAGGATTTGCAGTCATGATGTTCATACG
TCAGGCTAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCATGTGCCTGGGAATGTG
AGTGTTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATTCTAGTTTTGAATAATGTTCCCTAGTATACAGCAATTTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTCATATTCATAAGGAT
AGGTCTCATTCAACCACATGTTTATCATTTGGGATCTGCATTTTTCTGATGCAAAATGATTT
ATTCTTCCAGAGCACTGGAATTGGGTTGAATCATCTTATAACGGCCAAAACTAAATGCTTC
TGTGCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA
TGACAGAATGTCAGAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCCTTGGCAGGATTCTC
TTGGAACATCACACAAAGAAATTCCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACATGAGCTTTTAAAGTACGTTTGGCTGTTCAGACCTTGACTTTGAG
GTGAAGGAAAGCTCGCCAACTCCTTTTTATATGTAACACAATATATCAAGATCTAATGTGA
GACAGTATGCCAGTCCCAAGATCTGTCAATATGACTGAAGACACATTGGGATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCCAGTTCACTATAAAGCTTTATGAGAAATT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTTCTTCATTTACATTTC
AAATGCTATCCCAAAAGTCCCCTATACTCCCCCCTCACCCTGCTCCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCATATGATCCACCAAGCTTTACATGCTC
ACTATCTGGTCTATTGCAAGAATGGCTGCCGAGCTGATGCAGTCAGATACAGACACCTAC
AGCCAAACAGTGGAAGGAACTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATGG

GCCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
ACCCTTGGGGCTCTCAGAGTCTGAACCACCAACCATAGAACATTCATGGGCTGTACCCAG
GCCTCTCCACTCATATGTAACAGATATGTGGCTTGGCCTTCATCTGGGTCCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAGTTCTGAGGTGTCGGTGGAGGGTATACTCAGGGAGGCCCTCACCTGC
TAGGAAGAGAAGAGGAGGGGGAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAAATAAAAATAATTAAAAAAAAAAAAAGAATGGCTGCTGAGCCC
TACTCTAAAACCATTGCATCCCCCCCCCCCAATCATTCAGTGACTACGAATTAAAATCATT
GATACTAACAATAGATGTAGGAAACTATTGTTAACTTCTTTGTGACCACGAGTGGTATTTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTTCTTTCATTTCCCTGTACATGCATG
TAGCTTAATGATGTTCAGTGAATTAAAAATAACAATGAAGAATAAAGACAACTGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCACCTGGAAGAAATGTCTTCAGTT
TTTCTATATATGTTTACTCTATCGTATGTTAATTAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACTTATAAGAAAGTAAAACATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTTGAA
AGAATGTAATAACCTACTAAGGCAAACATAGAAGTTGAAATTAAAAAGAGTAGACAGGA
GAAGTAATAAGGCAAATAATGAATATTTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAATAATAAACATTTCCTACTTGAAAATTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

Figure 49

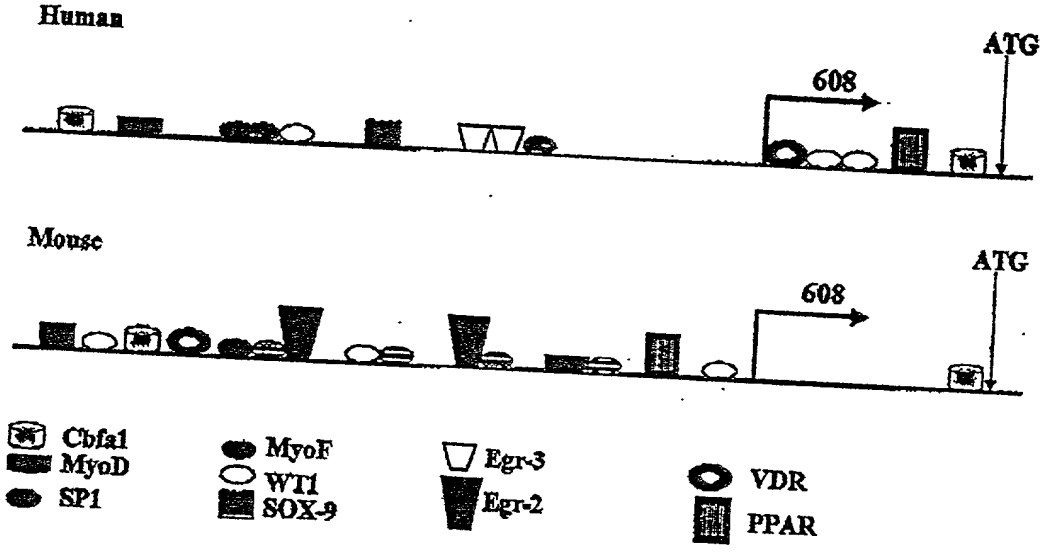

Human
ATG
608
Mouse
ATG
608
Cbfa1
MyoD
SP1
MyoF
WT1
SOX-9
Egr-3
Egr-2
VDR
PPAR

Figure 50

TAAGCCTTTTGCTCCCGTTGGAAGCAAAGAACGTTCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACTCTTTTGTTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAACCAACCCAGATGCCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTAAAAACAACGACTTTCATAAAATTTTTA
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAAACTGAGTCTTAAGGATTCATATTAACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTAACTTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTACTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTCGTGCCACACAGGTCTCCAGGG
TTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCATTCTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTTCTT
CACAAGTCTTTTTAGAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
TAAACTGACTTTAAAAAAAAAAGCTTATGATACTTGTCACAGAGTAAATCTTCCATA
AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTAAAAACCTTTAGAAA
TTTTGTTAACACAAAGATCCCTAGGCCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
TGAGAGATGGAGCTTAAGAACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
CAACAAACAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGTTTTTCCTTAATAGTCAA
ATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
GTAGATTAGAGGACACAAAGCAACAACTACAGAGTGAAGTAATCCAGAGACTTAAT
GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTTTT
ATCTTCATGTAGCTGATAATTTCATATTGTACACCTCAAACATAGATAACCAACAAA
GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTCGAAAGTCAGACA
CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
CTTGACTCTAGCTGCATATATATCAAAAGATGGCCTAATCGGCCATCACTGGAAAGA
GAGGCCATTGGACTTGCAAACTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAAA
AAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAAACTTA
CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
ACCACTACCAGCCCCAGCCCCCTTTGCCCCGCCCCCAACACACACACACACACACAC
ACACACACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCCTTCCCTGCAATAGTACCTC
TCCTATTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGG
GTTGGCGCTGGTGGTCCCGGGTCGATGGCCCGCCCATTCCCAGAAGCCGAGGCTATA
GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTTCAGGG
CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAAAATAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTCCTTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATG
TGGTATTCATTTACGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTTAAACTCATTCATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGGCCATGCTCCCCATCACTGTGTCAGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTATTTTGGAAGGTTAAGCCA
AATAAATAAAACACATATTAAATAATACAATGTTACAAAAATTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAATATTTTTAAAATG
TATATAAAAGTGTAAAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTTGTTGT
CATCATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAAACAAAAAACAAACCAACCAACCAAAAACCTCTGAAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTCATTTTTTTTCTTACTGTTTTGTTGTTGTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAAGGTTCATTTT
GCCACAAAAAGGTGTGAATTAAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAACTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCCACAGAGGTGCACTGTACATTTCGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCCTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTCACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACTTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTTCTTAATTAGTGATTCATGGG
GCAGGGCCCAGTCCATTGTTCGTGGTACCATTTCTCAGGCACTATTAAAAAAAAAAA
AACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCCTACCCTATTTGAGTTCCTGTCCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTTCCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTCACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGATGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTCAGAATCCCAGAGACT
CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACTTGAC
TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
GCATGAATAAGGGATCCCAGAGAGAAGGGAAGAAAAAGGAAGGAAGGAAGGAAG
GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAGAGGGAGGAAAGGAGGG
AGGGAAGGAAGGAAAGGGAAAGGAAAAAAGAGATGGGGAGGGAGGGAAGGAAAG
GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAAATAAATAAATTTTCAGGGATTATT
ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACTGAAACTACAATTATT
CCCACAAATGAAGAAAATGCTGTAAGAGTTTTCACACACCGAAGTGGGAAACTTAA
GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
CTCAACTCTTTAAGCAAACAATGCTGTTGTTCATGAAAAGCACAATAAAGTACATAT
GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
CACCCACTGTTCACAGAGACTTTAAACATGAAACTGGAACTATGTCTAGTGTTTTGA
CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTCATGTATCCTGCATGAAATTC
TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTG
AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTTATTAGTATTTCATGTTCTTT
TCTTTTTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGA
AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTT
TTTGCTTCTAATTCAGAATCAATTAAATTCATATTACTATAAAAGACAGCACGCCAG
ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCGGCTACATG
ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
TTAAAAAAAAAAAAAACTATTGATGTGGAATGATCTTAATTTGGGGAAAAGACAATT
TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAAATG
AGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT
GGAGACACCATAATAAAACTCCTTTTTTGTTTGCTAACTAAAACCACTAAAATCTAA
AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTA
CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAACAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTACTT
GTATGATAACTTCATTGACCTCCCTCCCAAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACTATCTTGTTTGTTTGTTTCTTTTTTT
ATARKACGTATTTTCCTCAATTTCATTTAGAATGATATCCCAAAAGTCCCCCATAACC
TCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCGTGTCCAATGGACCTCTCTTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

Figure 51

MPKRAHWGALSVVLILLWGHPRVALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
RINLGFNSIQALSETSFAGLTKLELLMIHGNEIPSIPDGALRDLSSLQVFKFSYNKLRVITG
QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
LSTIRHLYLAENMVRTLPASMLRNMPLLENLYLQGNPWTCDCEMRWFLEWDAKSRGIL
KCKKDKAYEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSIEEEQEQ
EEDGGSQLILEKFQLPQWSISLNMTDEHGNMVNLVCDIKKPMDVYKIHLNQTDPPDIDIN
ATVALDFECPMTRENYEKLWKLIAYYSEVPVKLHRELMLSKDPRVSYQYRQDADEEAL
YYTGVRAQILAEPEWVMQPSIDIQLNRRQSTAKKVLLSYYTQYSQTISTKDTRQARGRS
WVMIEPSGAVQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPDSKFSIL
SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDTVTIGKNPGES
VTLPCNALAIPEAHLSWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV
NQQGADHFTVGITVTKKGSGLPSKRGRRPGAKALSRVREDIVEDEGGSGMGDEENTSR
RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTPPFPAVSPPSASP
VQTVTSAEESSADVPLLGEEEHVLGTISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
SEKTEEITSTEGDLKGTAAPTLISEPYEPSPTLHTLDTVYEKPTHEETATEGWSAADVGSS
PEPTSSEYEPPLDAVSLAESEPMQYFDPDLETKSQPDEDKMKEDTFAHLTPTPTIWVNDS
STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQDTLLIKKGMKEMSQTLQGGNM
LEGDPTHSRSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETTVGTLLDKDTTTVTTTPR
QKVAPSSTMSTHPSRRRPNGRRRLRPNKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
QVESSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGTPRRKHGKRPNKHRYTPSTVSSRA
SGSKPSPSPENKHRNIVTPSSETILLPRTVSLKTEGPYDSLDYMTTTRKIYSSYPKVQETLP
VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
GTPTWNPSRTAQPGRLQTDIPVTTSGENLTDPPLLKELEDVDFTSEFLSSLTVSTPFHQEE
AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVAILLSETRPQNHTPTAARMKEPASSSPS
TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEGTQHMSGPNE
LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFTNKTLSF
PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPPLLHTPQTTGSPSTN
LQNIPMVSSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
TAETDTVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTSPRRTDVRYG
GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVVKSVTDK

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVFYGGDLKVDCVATGL
PNPEISWSLPDGSLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMRVRVKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKTVWIHVNVQPPKING
NPNPITTVREIAAGGSRKLIDCKAEGIPTPRVLWAFPEGVVLPAPYYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNCS
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNAA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPCTPPGAGQGRFSWTLPNGMHLEG
PQTLGRVSLLDNGTLTVREASVFDRGTYVCRMETEYGPSVTSIPVIVIAYPPRITSEPTPVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGSLTIQHAT
QRDAGFYKCMAKNILGSDSKTTYIHVF (SEQ ID NO: 21)

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCCGATGG AGCTTTAAGA
GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTTC AGACTCTCCA
CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
TCCGTGGACC TGCGATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
AATCCAGAGG AATTCTGAAG TGTAAAAAGG ACAAAGCTTA TGAAGGCGGT
CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAACTGC CCCAGTGGAG
CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
TCCAATGACC CGAGAAAACT ATGAAAAGCT ATGGAAATTG ATAGCATACT
ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA
CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
GTGCTACTTT CCTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT
GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTACTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
TGGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
AGACGCCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA
GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
GAAGCATTCG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
GGGCACAGAA GTACCCCCGT TGATTAAAAC CACAAGTCCT CCATCCTTGA
GCCTAGAAGT CACACCACCT TTTCCTGCTG TTTCTCCCCC CTCAGCATCT
CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCCTCA GCCAGCATGG
GGCTAGAACA CAACCACAAT GGAGTTATTC TTGTTGAACC TGAAGTAACA
AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
AGATGTTGGA TCGTCACCAG AGCCCACATC CAGTGAGTAT GAGCCTCCAT
TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTTAAT GACTCCAGTA
CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCCCAGGC
CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
TCGAAGGAGA CCCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
ACCGGCACAA GCAAACCCCA CCCACAACTT TTGCCCCATC AGAGACTTTT
TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
GAGTTCTCTG GTTCCTACAG CTTGGGTGGA TAACACAGTT AATACCCCCA
AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
AACATAGAAA CATTGTTACT CCCAGTTCAG AAACTATACT TTTGCCTAGA
ACTGTTTCTC TGAAAACTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
TTTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
AGGAAGAAGC TGGTTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
CTGCTGCCCG GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTC TTGAATTATG
TGGGGAATCC AGAAACAGAA GCAACCCCAG TCAACAATGA AGGAACACAG
CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATTC
CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTACTCCAAA
GTGTTTGGAA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCCAG
ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
TTCCTACAAC AGGATACATT CCCATAGCAC CTTCCATCTG GACTTTGGCC
CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCCTCA
ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
AGGATACAAC GGTTTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
GGTTCAAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTTCCT
GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
AGGCGTCCTT CTCAGACAGA GGCGTCTATA AGTGCGTGGC CAGCAATGCA
GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
GCATTCACAT TCACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
CCAAGGACAG CGGGCGCTAT GAGTGCGTGG CCGCCAACCT GGTAGGCTCC
GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
AAGATGCCGG AGATTACCTG TGCGTAGCTC GAAATAAGGT TGGTGATGAC
TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
GACGGGAGTC TGGTGAACTC CTTCATGCAG TCGGATGACA GCGGTGGACG
CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTTAACGAAG
TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
GGGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
ATACCAAGAT GGCACTCTCC TTATTCAGAA AGCCCAGCGT TCTGACAGCG
GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
GTGTGGATTC ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
CCATGGCAAC GGTTCCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC
TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCCAGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGTA
AGGGAAGCCA GGTTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTCATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCCTCTG TCACTTCAAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTTCTTCC
TGGAACTCAC TTTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAAAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTTC TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCCTT GTATGCAAAG TTTTTGTTGA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Figure 53

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGCCTGGTCGCCACC
CCTGGGGGCAAGGCCTGTCCTCGCCGCTGTGCCTGTTATATGCCTACGGAGGTACACTGCACATTTCGG
TACCTGACTTCCATCCCAGACAGCATCCCGCCCAATGTGGAACGCATCAATTTAGGATACAACAGCTTG
GTTAGATTGATGGAAACAGATTTTTCTGGCCTGACCAAACTGGAGTTACTCATGCTTCACAGCAATGGC
ATTCACACAATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAAT
AAAGTCCGAAAACTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGATTGCACATGGACCAC
AACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTCAACTTTCTCCGCCTGGTGCACTTGGAA
GGAAATCAGCTCACTAAGCTCCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAAATC
TCTTTCATTAAGTTCCTATACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTAT
ATGCCTGACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAGTGGTTG
TCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGCTCAG
CAGTGTCCACTTTGCATGAACCCTAGGACTTCTAAAGGCAAGCCGTTAGCTATGGTCTCAGCTGCAGCT
TTCCAGTGTGCCAAGCCAACCATTGACTCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGT
AGTTCTGCTTTCATCTCTCCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGAT
CAGTCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCCATTGCATTC
ACTGAAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTTGGTGTGCAACATAGATTAC
GGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTGTACAGTGATTCTCCTCTGATACTAGAAAGGAGC
CACTTGCTTAGTGAAACACCGCAGCTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATT
TTTACCAACATAGAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCCTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCTCAAATCACTTTA
CCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTG
GAACATACTGTCTTGGTAGGTGGAACCGTTGGCCTGAACTGCCCAGGCCAAGGAGACCCCACCCCACAC
GTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCCTTATGTCAGTGAGGATGGACGGATCCTA
ATAGACAAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTGACACAGGCGTATATCACTGTATA
AGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAACCTTTGGTCGAAGCC
TATCAGGAAAATGGGATTCATCACACAGTTTTCATTGGTGAAACACTTGATCTTCCATGCCATTCTACT
GGTATCCCAGATGCCTCTATTAGCTGGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGAC
AAGAAAGTTCTAAACAATGGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGC
TGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAAGGACAA
AGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCTT
AAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTGATGGAGGCTGAGGTTGGAAAACACACCTCA
AGCACAAGTAAGAGGCACAACTATCGGGAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGT
TTTAGGGAGAATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCACTG
TTGGAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCACCCCCAGTG
GTCACCCAACTCCCAAACATACCTGGTGAAGAAGACGATTCCTCAGGCATGCTCGCTCTACATGAGGAA
TTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAACAATA
TCTGATAGTCCTATGACAAACATAAATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTA
CCACCTGAAGAACCCACAGATTTCAAACTGTCTACTGCTATTAAAACTACAGCCATGTCAAAGAATATA
AACCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTCATCCACTGTCTTTCCACTGCTA
CTTGGAGCAACTGAATTTCAGGACTCTGACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCC
CCAATAACAGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGCACCACCAACAAACTA
TTATTAGAGTCAGTAAATACCACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCAC
AATCACTTCTATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCT
GCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTCAGACGCTTTGGGAGG
CAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATATAGAACTCCAGTTCTGCGACGGCATAGA
TACAGCATTTTCAGGTCAACAACCAGAGGTTCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTG
CTCAATGTGACATGTCTGTCCTGTCTTCCCAGGGAGAGGCTCACCACTGCCACAGCAGCATTGTCTTTT
CCAAGTGCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTACAACT
CTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAAACAACACCCACAATAAAA
TATTTCAGGACTGAAATTTCCCAAGTGACTCCAACTGGTGCAGTCATGACATATGCTCCAACATCCATA
CCCATGGAAAAAACTCACAAAGTAAACGCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGA
GATTCAGTGATTACATCGTCACTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACA
AGGTTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAAGGCAGATTA
AGGAATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCCTAAAACATCTCCTGCTTTA
CCACAGAGACAAAGTTCCCCTTTCCATTTCACCACACTTTCAACAAGTGTGATGCAAATTCCATCTAAT
ACCTTGACTACCGCTCACCACACTACGACCAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAG
CTTCCCTTCCCACCCCTTAACCCTATGCTTCCTAGTATTATAAGCAAAGACTCAAGTACAAAAAGCATC
ATATCAACGCAAACAGCAATACCAGCAACAACTCCTACCTTCCCTGCATCTGTCATCACTTATGAAACC
CAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAGCCTCAAAAGAAGAACAGGACTGAC
CCAAACATCTCTCCAGACCAGAGTTCTGGCTTCACTACACCCACTGCTATGACACCTCCTGCTCTGGCA

TTCACTCATTCCCCACCAGAAAACACAACTGGGATTTCAAGCACAATCAGTTTTCATTCAAGAACTCTT
AATCTGACAGATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTCT
GAAACAACTTTGTCCAGCAAATCACACCAGAGTACCACAACTAGGAAAGCATCATTAGACACTCCCATA
CCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCCATCTCCCCTCCCTTTACTCAGAGAGCA
GTTACTGACACACGTGGCGACTCCCATTTCCGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCC
TCAAGGCACAATCTCCAAATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTA
CATTCTACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACTCCATCTCCCTGG
GCAGAATACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAAGGCAAAAAGCCAGAAGTAAGC
ATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCTTGTTTCAGATTGGGATGGACAGAAGAACACA
AAGAAGAGTGACTTTGATAAGAAACCAGTTCAAGAAGCAACAACTTCCAAACTCCTTCCCTTTGACTCT
TTGTCTAGGTATATATTTGAAAAGCCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCT
AACTCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCCACCATTCATTGGACCAGAGTT
TCAGGACTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCTGTCC
ATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCGCATCCAATCTGTTTGGCACAGAC
CACCTTCATGTCACCTTGTCTGTGGTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATC
ACAGTTCATTCCGGAAGCACTGTGGAACTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCTGTGGTGACGGTT
GACGGAACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAAC
CCAGGTGGCCAGGATTCACTGCTGGTTAAAATACAAGTCATTGCAGCACCACCTGTTATTCTAGAGCAA
AGGAGGCAAGTCATTGTAGGCACTTGGGGTGAAAGTTTAAAACTGCCCTGTACTGCAAAAGGAACTCCT
CAGCCCAGCGTTTACTGGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCAAG
TTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGACAGGGGCACTTATGAA
TGCATTGCTACCAGTTCCACTGGTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGACC
AGCCCCAGGATAGAAGCTGCATCCCAGAAAAGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAAC
TGCTCAGCCACTGGGGAGCCCAAACCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTCGACCAG
TGGAGCTGGATCCACGTCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGT
GTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCATGTTAGCCTAAGACTG
AAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTA
GATTGCAAAGCTTCCGGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAAC
AATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACTTTA
TACTTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGCTATGCCCAGAACACCCTAGGGAAA
GATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATAAGGCAGAGTAACAAAACCAAC
AAGAGAATCAAAGCTGGAGACACAGCTGTCCTTGACTGTGAGGTCACTGGGGATCCCAAACCAAAAATA
TTTTGGTTGCTGCCTTCCAATGACATGATTTCCTTCTCCATTGATAGGTACACATTTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAAACTGGATGTGGTCTCTAAACCTCCATTAATCAATGGTCTGTATACA
AACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCTGAAGGG
ACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTTCCTCACAGCCCCATACTATGGAAGC
AGAATCACAGTCCATAAAAATGGAACCTTGGAAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTT
ATCTGTGTGGCCCGAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAAATGCTG
AGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGAAAGTCCACAGCATTG
AATTGCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACACGATTTTCC
AATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAACAACTCGG
GAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTA
GAAATTGGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAGAATCT
CTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGGACTATGCCAAGTGGTTAT
GTAGTAGACAGGCCTCAAATTAATGGGAAATACATATTGCATGACAATGGCACCTTAGTCATTAAAGAA
GCAACAGCTTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTTGGTCATACACTGATTACT
GTTCCAGTAATGATTGTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGG
ACAGGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACATGGAGATG
CCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGAAGTGAGCAGCTTCACTTACAA
GGTACCCTAGTCATTCAGAATCCCCAAACCTCCGATTCTGGGATATACAAATGCACAGCAAAGAACCCA
CTTGGTAGTGATTATGCAGCAACGTATATTCAAGTAATCTGA

Figure 54

ORIGIN

```
1       MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYMPTE VHCTFRYLTS
51      IPDSIPPNVE RINLGYNSLV RLMETDFSGL TKLELLMLHS NGIHTIPDKT
101     FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
151     GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
201     PQEMVSYMPD LDSLYLHGNP WTCDCHLKWL SDWIQEKPDV IKCKKDRSPS
251     SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
301     SAFISPQGFM APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
351     YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
401     YYKYKQVAPK PEDIFTNIEA DLRADPSWLM QDQISLQLNR TATTFSTLQI
451     QYSSDAQITL PRAEMRPVKH KWTMISRDNN TKLEHTVLVG GTVGLNCPGQ
501     GDPTPHVDWL LADGSKVRAP YVSEDGRILI DKSGKLELQM ADSFDTGVYH
551     CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
601     PDASISWVIP GNNVLYQSSR DKKVLNNGTL RILQVTPKDQ GYYRCVAANP
651     SGVDFLIFQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
701     SALMEAEVGK HTSSTSKRHN YRELTLQRRG DSTHRRFREN RRHFPPSARR
751     IDPQHWAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
801     ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVNSQI
851     LPPEEPTDFK LSTAIKTTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
901     TEFQDSDQMG RGREHFQSRP PITVRTMIKD VNVKMLSSTT NKLLLESVNT
951     TNSHQTSVRE VSEPRHNHFY SHTTQILSTS TFPSDPHTAA HSQFPIPRNS
1001    TVNIPLFRRF GRQRKIGGRG RIISPYRTPV LRRHRYSIFR STTRGSSEKS
1051    TTAFSATVLN VTCLSCLPRE RLTTATAALS FPSAAPITFP KADIARVPSE
1101    ESTTLVQNPL LLLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTSI
1151    PMEKTHKVNA SYPRVSSTNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
1201    RKIPWQQNFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
1251    HFTTLSTSVM QIPSNTLTTA HHTTTKTHNP GSLPTKKELP FPPLNPMLPS
1301    IISKDSSTKS IISTQTAIPA TTPTFPASVI TYETQTERSR AQTIQREQEP
1351    QKKNRTDPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
1401    RTLNLTDVIE ELAQASTQTL KSTIASETTL SSKSHQSTTT RKASLDTPIP
1451    PFLSSSATLM PVPISPPFTQ RAVTDTRGDS HFRLMTNTVV KLHESSRHNL
1501    QMPSSQLEPL TSSTSNLLHS TPMPALTTVK SQNSKLTPSP WAEYQFWHKP
1551    YSDIAEKGKK PEVSMLATTG LSEATTLVSD WDGQKNTKKS DFDKKPVQEA
1601    TTSKLLPFDS LSRYIFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
1651    IHWTRVSGLD LSRGNQNSRV QVLPNGTLSI QRVEIQDRGQ YLCSASNLFG
1701    TDHLHVTLSV VSYPPRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
1751    LANQTVVSES SQGSRQAVVT VDGTLVLNNL SYYDRGFYKC VASNPGGQDS
1801    LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYWVLSD
1851    GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRVV
1901    MLTMEERVTS PRIEAASQKR TEVNFGDKLL LNCSATGEPK PQIMWRLPSK
1951    AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNKMGD DLILMHVSLR
2001    LKPAKIDHKQ YFRKQVLHGK DFQVDCKASG SPVPEISWSL PDGTMINNAM
2051    QADDSGHRTR RYTLFNNGTL YFNKVGVAEE GDYTCYAQNT LGKDEMKVHL
2101    TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
2151    FSIDRYTFHA NGSLTINKVK LLDSGEYVCV ARNPSGDDTK MYKLDVVSKP
2201    PLINGLYTNR TVIKATAVRH SKKHFDCRAE GTPSPEVMWI MPDNIFLTAP
2251    YYGSRITVHK NGTLEIRNVR LSDSADFICV ARNEGGESVL VVQLEVLEML
2301    RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
2351    SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPVI
2401    LTYAPGTVKG ISGESLSLHC VSDGIPKPNI KWTMPSGYVV DRPQINGKYI
2451    LHDNGTLVIK EATAYDRGNY ICKAQNSVGH TLITVPVMIV AYPPRITNRP
2501    PRSIVTRTGA AFQLHCVALG VPKPEITWEM PDHSLLSTAS KERTHGSEQL
2551    HLQGTLVIQN PQTSDSGIYK CTAKNPLGSD YAATYIQVI*
```

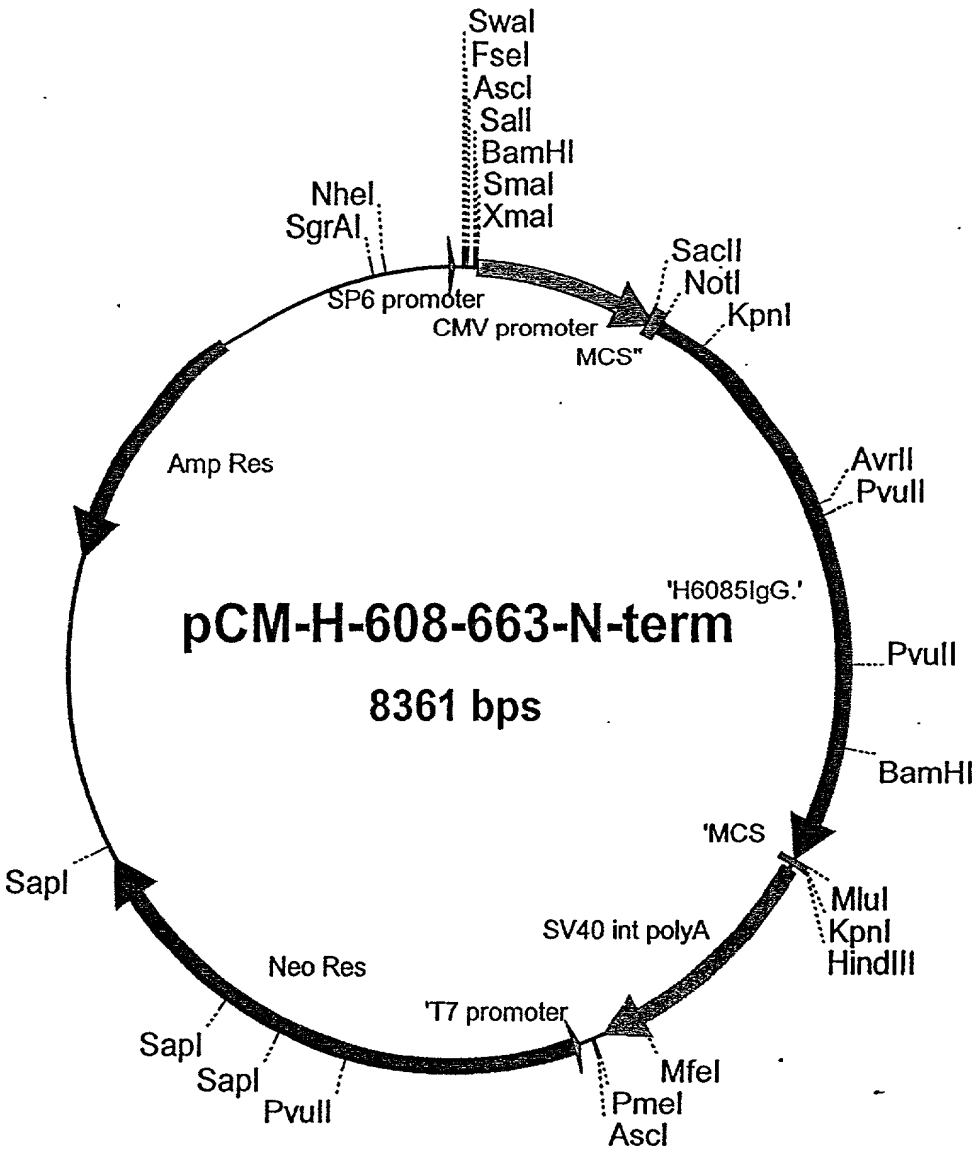
Figure 55
OCP rat amino acid sequence

>608-663Nterm Rat Protein (663 aa)

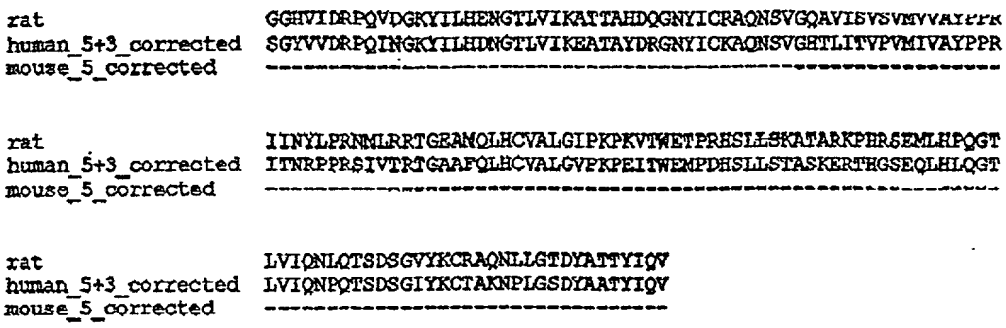
MQVRGREVSGLLISLTAVCLVVTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
RINLGYNSLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
IFKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESLYLHGNPWTCDCHLKWLSEWMQGNPDI
IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSKSLVTQEDNG
SASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFST
NLVCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEA
DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNN
PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQM
ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDLPCLSTGV
PDASISWILPGNTVFSQPSRDRQILNNGTLRILQVTPKDQGHYQCVAANPSGADFSSFKV
SVQ

Figure 56

pCM-H-608-663-N-term



```
rat                     -TSKHLPYVSLPKTLLKKPRIIGGKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGK
human_5+3_corrected     TTSKLLPFDSLSRYIFEKPRIVGGKAASFTIPANSDAFLPCEAVGNPLPTIHWTRVS-GL
mouse_5_corrected       ------------------------------------------------------------

rat                     KISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILD
human_5+3_corrected     DLSRGNQNSRVQVLPNGTLSIQRVEIQDRGQYLCSASNLFGTDHLHVTLSVVSYPPRILE
mouse_5_corrected       ------------------------------------------------------------

rat                     RHVKEITVHFGSTVELKCRVEGMPRPTVSWILANQTVVSETAKGSRKVWVTPDGTLIIYN
human_5+3_corrected     RRTKEITVHSGSTVELKCRAEGRPSPTVTWILANQTVVSESSQGSRQAVVTVDGTLVLHN
mouse_5_corrected       ------------------------------------------------------------

rat                     LSLYDRGFYKCVASNPSGQDSLLVKIQVITAPPVIIEQKRQAIVGVLGGSLKLPCTAKGT
human_5+3_corrected     LSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRRQVIVGTWGESLKLPCTAKGT
mouse_5_corrected       ------------------------------------------------------------

rat                     PQPSVHWVLYDGTELKPLQLTHSRFFLYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV
human_5+3_corrected     PQPSVYWVLSDGTEVKPLQFTNSKLFLFSNGTLYIRNLASSDRGTYECIATSSTGSERRV
mouse_5_corrected       ------------------------------------------------------------

rat                     VILTVEEGETIPRIETASQKWTEVNLGEKLLLNCSATGDPKPRIIWRLPSKAVIDQWHRM
human_5+3_corrected     VMLTMEERVTSPRIEAASQKRTEVNFGDKLLLNCSATGEPKPQIMWRLPSKAVVDQ----
mouse_5_corrected       ------------------------------------------------------------

rat                     GSRIHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLVLMHVRLRLTPAKIEQKQYFKKQ
human_5+3_corrected     GSWIHVYPNGSLFIGSVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDHKQYFRKQ
mouse_5_corrected       ------------------------------------------------------------

rat                     VLHGKDFQVDCKASGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFNNV
human_5+3_corrected     VLHGKDFQVDCKASGSPVPEISWSLPDGTNINNAMQADDSGHRTRRYTLFNNGTLYFNKV
mouse_5_corrected       ------------------------------------------------------------

rat                     GMAEEGDYICSAQNTLGKDEMKVHLTVLTAIPRIRQSYKTTMRLRAGETAVLDCEVTGEP
human_5+3_corrected     GVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTNKRIKAGDTAVLDCEVTGDP
mouse_5_corrected       ------------------------------------------------------------

rat                     KPNVFWLLPSNNVISFSNDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLD
human_5+3_corrected     KPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSGDDTKMYKLD
mouse_5_corrected       ------------------------------------------------------------

rat                     IVSKPPLINGLYANKTVIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSR
human_5+3_corrected     VVSKPPLINGLYTNRTVIKATAVRHSKKHFDCRAEGTPSPEVMWIMPDNIFLTAPYYGSR
mouse_5_corrected       ------------------------------------------------------------

rat                     VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIA
human_5+3_corrected     ITVHKNGTLEIRNVRLSDSADFICVARNEGGESVLVVQLEVLEMLRRPTFRNPFNEKIVA
mouse_5_corrected       ------------------------------------------------------------

rat                     QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPHNSPYILMAGNGSLILYKATRNKSGKY
human_5+3_corrected     QLGKSTALNCSVDGNPPPEIIWILPNGTRFSNGPQSYQYLIASNGSFIISKTTREDAGKY
mouse_5_corrected       ------------------------------------------------------------

rat                     RCAARNKVGYIEKLILLEIGQKPVILTYRPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTP
human_5+3_corrected     RCAARNKVGYIEKLVILEIGQKPVILTYAPGTVKGISGESLSLHCVSDGIPKPNIKWTMP
mouse_5_corrected       ------------------------------------------------------------
```



(rat: SEQ ID NO:10)
(human_5+3_corrected: SEQ ID NO:11)
(mouse_5_corrected: SEQ ID NO:12)